计算数学丛书

对称矩阵计算

蒋尔雄 编著

上海科学技术出版社

计算数学丛书

对称矩阵计算

蒋 尔 雄 编著

上海科学技术出版社

计算数学丛书

对称矩阵计算

蒋 尔 雄 编著

上海科学技术出版社出版
(上海瑞金二路 450 号)

新华书店上海发行所发行　上海商务印刷厂印刷

开本 787×1092 1/32　印张 5.375　字数 115,000
1984 年11月第1版　1984 年11月第1次印刷
印数：1—11,500

统一书号：13119·1184　　定价：0.75 元

出版说明

《计算数学丛书》是为了适应计算数学和计算机科学的发展，配合高等院校计算数学教学的需要而组织的一套参考读物。读者对象主要是高等院校数学系和计算机科学系的学生、研究生，亦可供高等院校数学系和计算机科学系的教师以及工矿企业、科研单位从事计算工作的技术人员参考。

本丛书向读者介绍近代计算方法的一些主要进展及其适用范围和实用效果。每种书集中介绍一个专题，针对本专题的近代发展作综合性的介绍，内容简明扼要，重点突出，有分析，有评价，力图使读者对该专题的动向和发展趋势得到一个完整的了解。

本丛书已拟定的选题计有：《线性代数与多项式的快速算法》、《数论变换》、《数值有理逼近》、《矩阵特征值问题》、《索伯列夫空间引论》、《计算组合数学》、《样条与插值》、《有限条形法》、《广义逆矩阵及其计算方法》、《非线性方程迭代解法》、《奇异摄动中的边界层校正法》、《沃尔什函数理论与应用》、《多项式最佳逼近的实现》、《坏条件常微分方程数值解》、《误差分析》、《最小二乘问题的数值解法》、《板壳问题非协调方法》、《外推法及其应用》、《Monte Carlo 方法》、《差分格式理论》、《高维偏微分方程数值解》等二十余种，于一九八〇年初起陆续出版。

引 言

矩阵计算是近代计算方法中发展最早、最快的领域. 尤其是对称矩阵计算是最成熟的部分.

本书比较深入地介绍了解线性代数方程组的共轭斜量法和 Lanczos 方法, 解矩阵特征值问题的 QL 方法和 Lanczos 方法, 以及与这些方法密切相关的对称三对角矩阵的理论. 这些方法是经过实践考验的, 认为是有效的方法. 其中解线性代数方程组的共轭斜量法和解特征值问题的 Lanczos 方法, 都是五十年代初期就提出来了, 经历了二十年曲折的历史, 获得了理论上和实践上的支持, 直到七十年代才站住了脚.

读者通过本书, 一方面可以更好地掌握这些方法, 应用于实际, 也可以了解到目前在对称矩阵计算理论上达到的深度; 并且会发现, 尽管对称矩阵计算这个领域, 在计算方法各个领域中, 相对来说是比较成熟的, 但仍然有不少问题有待解决.

本书的初稿曾是复旦大学计算数学专业 77 届、78 届学生选修课的讲稿, 在教学过程中, 发现学生对这些内容还是比较感兴趣的. 现在写成此书, 作为大学计算数学专业大学生的选修课和研究生课程的教材也是适当的.

本书中有很多素材来自美国 California 大学 Berkeley 分校数学系和计算机科学系教授 B. N. Parlett 的书 "The Symmetric Eigenvalue Problem", 记得在 1979 年那书正式出版之前, Parlett 教授把这书的打印稿给我阅读, 使我能及时掌握有关的理论. 在此向 Parlett 教授表示衷心的感谢.

本书的其它材料，来自书后所列的参考文献和自己的心得．由于水平有限，错误和不妥之处在所难免，望读者批评指正．

蒋尔雄

于复旦大学　1982年10月

记号说明

本书规定

1. 用大写英文字母表示矩阵，用黑体小写英文字母表示向量，用小写英文字母或希腊字母表示标量.

2. A^T, $\boldsymbol{x}^T$ 表示 A, $\boldsymbol{x}$ 的转置; A^*, $\boldsymbol{x}^*$ 表示 A, $\boldsymbol{x}$ 的转置共轭.

3. 列向量 $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)^T$, $\boldsymbol{y}=(y_1, y_2, \cdots, y_n)^T$ 的内积

$$(\boldsymbol{x}, \boldsymbol{y})=\sum_{i=1}^{n}\bar{x}_i y_i=\boldsymbol{x}^*\boldsymbol{y}.$$

4. 向量 $\boldsymbol{x}$ 的范数 $\|\boldsymbol{x}\|$ 表示 $\boldsymbol{x}$ 的 l_2 范数, 矩阵 A 的范数

$$\|A\|=\sup_{\boldsymbol{x}\neq 0}\|A\boldsymbol{x}\|/\|\boldsymbol{x}\|.$$

5. $\boldsymbol{e}_i$ 表示单位矩阵

$$\begin{pmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots \\ 0 & & & 1 \end{pmatrix}$$

的第 i 列.

6. $\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j$ 是 j 个 n 维列向量, 记$[\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j]$表示由这 j 个向量, 按照所给的次序组成的矩阵.

$\{\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j\}$表示由这 j 个向量所成的线性空间.

7. 没有特殊声明，矩阵、向量、常数全是指实矩阵、实向量、实数；任意矩阵，意指任意实矩阵。

目　录

第 1 章

共轭斜量法

共轭斜量法 (conjugate gradient method) 是解系数矩阵为对称正定的线性代数方程组的一种方法，它的产生在五十年代初期，参见参考文献[1]；经过几十年的考验，现在公认为它是一种好方法，可详阅[2]、[3]、[4]. 为了深入了解共轭斜量法，我们从最佳逼近的观点来介绍共轭斜量法. 另外我们也介绍一下，近年来产生的，引起广泛注意的不完全分解、预处理、共轭斜量法.

§1 斜 量 法

设 $A=(a_{ij})$ 是 $n\times n$ 实对称、正定矩阵. 考虑线性方程组

$$A\boldsymbol{x}=\boldsymbol{b} \tag{1}$$

的求解问题. 这里 $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)^T$ 是未知向量，$\boldsymbol{b}=(b_1, b_2, \cdots, b_n)^T$ 是已知的向量.

(1) 的求解问题，等价于下列泛函的求极小问题：

$$F(\boldsymbol{x})=(A\boldsymbol{x}, \boldsymbol{x})-2(\boldsymbol{b}, \boldsymbol{x})$$

或

$$F(\boldsymbol{x})=\boldsymbol{x}^TA\boldsymbol{x}-2\boldsymbol{b}^T\boldsymbol{x}, \tag{2}$$

即使 (2) 达到极小的向量 $\tilde{\boldsymbol{x}}$ 为 (1) 的解，反之，(1) 的解是使 (2) 达到极小的向量.

实际上，因为 A 正定，故 A^{-1} 存在，记 $\tilde{\boldsymbol{x}}=A^{-1}\boldsymbol{b}$，它是 (1) 的解，由

$$
\begin{aligned}
F(\boldsymbol{x}) &= (A\boldsymbol{x},\ \boldsymbol{x}) - 2(\boldsymbol{b},\ \boldsymbol{x}) = (A\boldsymbol{x},\ \boldsymbol{x}) - 2(A\tilde{\boldsymbol{x}},\ \boldsymbol{x}) \\
&= (A(\boldsymbol{x}-\tilde{\boldsymbol{x}}),\ (\boldsymbol{x}-\tilde{\boldsymbol{x}})) - (A\tilde{\boldsymbol{x}},\ \tilde{\boldsymbol{x}}) \\
&\geqslant -(A\tilde{\boldsymbol{x}},\ \tilde{\boldsymbol{x}}) = F(\tilde{\boldsymbol{x}})
\end{aligned}
$$

且因为 A 是正定的，故只有当 $\boldsymbol{x}-\tilde{\boldsymbol{x}}=0$ 时，才能使上式中的 $\geqslant$ 成为等号，否则是 $>$ 号. 这就证明了(1)的求解问题，等价于(2)的求极小问题.

(2)的求极小问题，等价于

$$
F_1(\boldsymbol{x}) = (A(\boldsymbol{x}-\tilde{\boldsymbol{x}}),\ (\boldsymbol{x}-\tilde{\boldsymbol{x}})) \tag{3}
$$

或

$$
F_1(\boldsymbol{x}) = (A^{-1}(A\boldsymbol{x}-\boldsymbol{b}),\ (A\boldsymbol{x}-\boldsymbol{b})) \tag{4}
$$

的求极小问题. 这是因为 $F_1(\boldsymbol{x}) = F(\boldsymbol{x}) + (A\tilde{\boldsymbol{x}},\ \tilde{\boldsymbol{x}})$，而 $(A\tilde{\boldsymbol{x}},\ \tilde{\boldsymbol{x}})$ 是常数.

(3)或(4)与(2)的主要差别在于(2)表示式中没有未知量，因此给定一个 $\boldsymbol{x}$，可以算出 $F(\boldsymbol{x})$，而(3)，(4)都有未知量出现，当不知道 $\tilde{\boldsymbol{x}}$ 或 A^{-1} 时，给定一个 $\boldsymbol{x}$，算不出 $F_1(\boldsymbol{x})$. 但当不需要计算泛函值时，用 $F_1(\boldsymbol{x})$ 更加直观.

一种求(1)解的想法是先取一个初始向量 $\boldsymbol{x}_0$，按某种规则求出一个向量 $\boldsymbol{x}_1$，使得 $F_1(\boldsymbol{x}_1) < F_1(\boldsymbol{x}_0)$，然后再从 $\boldsymbol{x}_1$，按上述规则求出一个向量 $\boldsymbol{x}_2$，使得 $F_1(\boldsymbol{x}_2) < F_1(\boldsymbol{x}_1)$，依此类推，得到一个向量序列 $\boldsymbol{x}_0$、$\boldsymbol{x}_1$、$\boldsymbol{x}_2$、…，并且希望这个序列能收敛到解 $\tilde{\boldsymbol{x}}$. 这是一种很一般的想法，很多方法都是基于这样的想法. 当然所得到的向量序列能不能收敛到 $\tilde{\boldsymbol{x}}$，收敛速度如何，都依赖于按什么样的规则从 $\boldsymbol{x}_i$ 确定 $\boldsymbol{x}_{i+1}$.

斜量法也是一种具体实现这种想法的方法. 它的规则是：从 $\boldsymbol{x}_0$，有 $A\boldsymbol{x}_0-\boldsymbol{b}=\boldsymbol{r}_0$，如果 $\boldsymbol{r}_0=0$，那么 $\boldsymbol{x}_0$ 即为(1)的解；如果 $\boldsymbol{r}_0\neq 0$，那么令 $\boldsymbol{x}=\boldsymbol{x}_0+\alpha\boldsymbol{r}_0$，当 α 变动时，表示一条过 $\boldsymbol{x}_0$ 的直线，它的方向跟 $\boldsymbol{r}_0$ 相同. 在这条直线上找一点 $\boldsymbol{x}_1=$

$\boldsymbol{x}_0+\alpha_0\boldsymbol{r}_0$，使得对所有的实数 α，

$$F_1(\boldsymbol{x}_1)\leqslant F_1(\boldsymbol{x}_0+\alpha\boldsymbol{r}_0), \tag{5}$$

也即在这条直线上，$\boldsymbol{x}_1$ 使 $F_1(\boldsymbol{x})$ 达到极小.

现在来看 α_0 是什么？因为

$$\begin{aligned}F_1(\boldsymbol{x}_0+\alpha\boldsymbol{r}_0)&=(A(\boldsymbol{x}_0+\alpha\boldsymbol{r}_0-\tilde{\boldsymbol{x}}),\ (\boldsymbol{x}_0+\alpha\boldsymbol{r}_0-\tilde{\boldsymbol{x}}))\\&=(A(\boldsymbol{x}_0-\tilde{\boldsymbol{x}}),\ (\boldsymbol{x}_0-\tilde{\boldsymbol{x}}))\\&\quad+2\alpha(A(\boldsymbol{x}_0-\tilde{\boldsymbol{x}}),\ \boldsymbol{r}_0)+\alpha^2(A\boldsymbol{r}_0,\ \boldsymbol{r}_0),\end{aligned}$$

有 $$\frac{\partial F(\boldsymbol{x}_0+\alpha\boldsymbol{r}_0)}{\partial\alpha}=2(A(\boldsymbol{x}_0-\tilde{\boldsymbol{x}}),\ \boldsymbol{r}_0)+2\alpha(A\boldsymbol{r}_0,\ \boldsymbol{r}_0),$$

$$\frac{\partial^2F}{\partial\alpha^2}=2(A\boldsymbol{r}_0,\ \boldsymbol{r}_0)>0,$$

因此取 $\dfrac{\partial F_1}{\partial\alpha}=0$ 得到的 α 即为 α_0，即

$$\begin{aligned}\alpha_0&=-(A(\boldsymbol{x}_0-\tilde{\boldsymbol{x}}),\ \boldsymbol{r}_0)/(A\boldsymbol{r}_0,\ \boldsymbol{r}_0)\\&=-(\boldsymbol{r}_0,\ \boldsymbol{r}_0)/(A\boldsymbol{r}_0,\ \boldsymbol{r}_0),\end{aligned}$$

从而 $\boldsymbol{x}_1=\boldsymbol{x}_0-[(\boldsymbol{r}_0,\ \boldsymbol{r}_0)/(A\boldsymbol{r}_0,\ \boldsymbol{r}_0)]\boldsymbol{r}_0$，而

$$\begin{aligned}F_1(\boldsymbol{x}_1)&=(A(\boldsymbol{x}_0-\tilde{\boldsymbol{x}}),\ (\boldsymbol{x}_0-\tilde{\boldsymbol{x}}))\\&\quad-2[(\boldsymbol{r}_0,\ \boldsymbol{r}_0)/(A\boldsymbol{r}_0,\ \boldsymbol{r}_0)](\boldsymbol{r}_0,\ \boldsymbol{r}_0)\\&\quad+[(\boldsymbol{r}_0,\ \boldsymbol{r}_0)/(A\boldsymbol{r}_0,\ \boldsymbol{r}_0)]^2(A\boldsymbol{r}_0,\ \boldsymbol{r}_0)\\&=F(\boldsymbol{x}_0)-(\boldsymbol{r}_0,\ \boldsymbol{r}_0)^2/(A\boldsymbol{r}_0,\ \boldsymbol{r}_0)<F_1(\boldsymbol{x}_0).\end{aligned}$$

上面的 α_0 是通过 $\dfrac{\partial F_1}{\partial\alpha}=0$ 而导出的，也可通过另一种思想来导出. 对于 R^n 引进一种新的内积：$[\boldsymbol{x},\ \boldsymbol{y}]=(A\boldsymbol{x},\ \boldsymbol{y})$，因为 A 是对称正定的，因此它满足内积的条件（参见[9]）从而有一种新的范数 $|\boldsymbol{x}|=\sqrt{(A\boldsymbol{x},\ \boldsymbol{x})}=\sqrt{[\boldsymbol{x},\ \boldsymbol{x}]}$. 于是

$$F_1(\boldsymbol{x})=|\boldsymbol{x}-\tilde{\boldsymbol{x}}|^2,$$

$$F_1(\boldsymbol{x}_0+\alpha\boldsymbol{r}_2)=|\boldsymbol{x}_0-\tilde{\boldsymbol{x}}+\alpha\boldsymbol{r}_0|^2,$$

求 x_1 的问题，相当于向量 $\boldsymbol{x}_0-\tilde{\boldsymbol{x}}$ 减去 $\boldsymbol{r}_0$ 的某一倍数后，使

余量最小. 自然 α_0 使得 $\boldsymbol{x}_0-\tilde{\boldsymbol{x}}+\alpha_0\boldsymbol{r}_0$ 正交于 $\boldsymbol{r}_0$. 即 α_0 满足

$$[\boldsymbol{x}_0-\tilde{\boldsymbol{x}}+\alpha_0\boldsymbol{r}_0,\ \boldsymbol{r}_0]=0,$$

$$\alpha_0=-\frac{[\boldsymbol{x}_0-\tilde{\boldsymbol{x}},\ \boldsymbol{r}_0]}{[\boldsymbol{r}_0,\ \boldsymbol{r}_0]}=-\frac{(\boldsymbol{r}_0,\ \boldsymbol{r}_0)}{(A\boldsymbol{r}_0,\ \boldsymbol{r}_0)} \tag{6}$$

时,可使余量 $|\boldsymbol{x}_0+\alpha_0\boldsymbol{r}_0-\tilde{\boldsymbol{x}}|$ 最小.

这种由正交性导出极值性的办法是有普遍意义的.

定理 1.1 R^n 上定义的任意一种内积 $((\cdot,\ \cdot))$, $\boldsymbol{y}$ 是 R^n 中的任意一个向量, $\boldsymbol{f}_1$、$\boldsymbol{f}_2$、…、$\boldsymbol{f}_k$ 是 R^n 中 k 个线性无关的向量,则 k 个常数 α_1^0、α_2^0、…、α_k^0 使

$$((\boldsymbol{y}+\alpha_1^0\boldsymbol{f}_1+\cdots+\alpha_k^0\boldsymbol{f}_k,\ \boldsymbol{y}+\alpha_1^0\boldsymbol{f}_1+\cdots+\alpha_k^0\boldsymbol{f}_k))$$
$$<((\boldsymbol{y}+\alpha_1\boldsymbol{f}_1+\cdots+\alpha_k\boldsymbol{f}_k,\ \boldsymbol{y}+\alpha_1\boldsymbol{f}_1+\cdots+\alpha_k\boldsymbol{f}_k)), \tag{7}$$

对任意实数 α_1、α_2、…、α_k 只要 $(\alpha_1, \alpha_2, \cdots, \alpha_k)\neq(\alpha_1^0, \alpha_2^0, \cdots, \alpha_k^0)$, 成立的充分必要条件是: α_1^0、α_2^0、…、$\alpha_k^{(0)}$ 满足

$$((\boldsymbol{y}+\alpha_1^0\boldsymbol{f}_1+\cdots+\alpha_k^0\boldsymbol{f}_k,\ \boldsymbol{f}_i))=0,\quad i=1、2、\cdots、k. \tag{8}$$

证明 若(8)成立,记 $\alpha_i=\alpha_i^0+\Delta\alpha_i$, 则

$$\begin{aligned}
&\left(\left(\boldsymbol{y}+\sum_{i=1}^{k}\alpha_i\boldsymbol{f}_i,\ \boldsymbol{y}+\sum_{i=1}^{k}\alpha_i\boldsymbol{f}_i\right)\right)\\
&=\left(\left(\boldsymbol{y}+\sum_{i=1}^{k}\alpha_i^0\boldsymbol{f}_i+\sum_{i=1}^{k}\Delta\alpha_i\boldsymbol{f}_i,\ \boldsymbol{y}+\sum_{i=1}^{k}\alpha_i^0\boldsymbol{f}_i+\sum_{i=1}^{k}\Delta\alpha_i\boldsymbol{f}_i\right)\right)\\
&=\left(\left(\boldsymbol{y}+\sum_{i=1}^{k}\alpha_i^0\boldsymbol{f}_i,\ \boldsymbol{y}+\sum_{i=1}^{k}\alpha_i^0\boldsymbol{f}_i\right)\right)\\
&\quad+2\left(\left(\boldsymbol{y}+\sum_{i=1}^{k}\alpha_k^0\boldsymbol{f}_i,\ \sum_{i=1}^{k}\Delta\alpha_i\boldsymbol{f}_i\right)\right)\\
&\quad+\left(\left(\sum_{i=1}^{k}\Delta\alpha_i\boldsymbol{f}_i,\ \sum_{i=1}^{k}\Delta\alpha_i\boldsymbol{f}_i\right)\right)
\end{aligned}$$

中间项 $\left(\left(\boldsymbol{y}+\sum\alpha_i^0\boldsymbol{f}_i,\ \sum_{i=1}^{k}\Delta\alpha_i\boldsymbol{f}_i\right)\right)$ 为 0, 而第 3 项因为 $\boldsymbol{f}_1$、$\boldsymbol{f}_2$、…、$\boldsymbol{f}_k$ 线性无关,除非所有 $\Delta\alpha_i(i=1、2、\cdots、k)$ 为 0, 否则总是正的, 因此(7)成立.

反之若(7)成立，取

$$(\alpha_1, \alpha_2, \cdots, \alpha_k)=(\alpha_1^0, \cdots, \alpha_{i-1}^0, \alpha_i^0+\Delta\alpha_i, \alpha_{i+1}^0, \cdots, \alpha_k^0),$$

于是

$$\begin{aligned}&\left(\left(\boldsymbol{y}+\sum_{l=1}^{k}\alpha_l\boldsymbol{f}_l, \boldsymbol{y}+\sum_{l=1}^{k}\alpha_l\boldsymbol{f}_l\right)\right)\\&\quad=\left(\left(\boldsymbol{y}+\sum_{l=1}^{k}\alpha_l^0\boldsymbol{f}_l, \boldsymbol{y}+\sum_{l=1}^{k}\alpha_l^0\boldsymbol{f}_l\right)\right)\\&\qquad+2\Delta\alpha_i\left(\left(\boldsymbol{y}+\sum_{l=1}^{k}\alpha_l^0\boldsymbol{f}_l, \boldsymbol{f}_i\right)\right)+(\Delta\alpha_i)^2((\boldsymbol{f}_i, \boldsymbol{f}_i)),\end{aligned}$$

由(7)知 $\Delta\alpha_i=0$ 使上式达到极小，故必须 $2\Delta\alpha_i$ 的系数

$$\left(\left(\boldsymbol{y}+\sum_{l=1}^{k}\alpha_l^0\boldsymbol{f}_l, \boldsymbol{f}_i\right)\right)=0,$$

否则因为 $((\boldsymbol{f}_i, \boldsymbol{f}_i))>0$，取

$$\Delta\alpha_i=-\left(\left(\boldsymbol{y}+\sum_{l=1}^{k}\alpha_l^0\boldsymbol{f}_k, \boldsymbol{f}_i\right)\right)\Big/((\boldsymbol{f}_i, \boldsymbol{f}_i))\neq 0,$$

将使该式达到极小，跟假设(7)成立矛盾.

因为所取的 i 是任意的，因此(8)成立. 定理证毕.

再回到方程(1)的求解问题上来. 求得了 $\boldsymbol{x}_1=\boldsymbol{x}_0+\alpha_0\boldsymbol{r}_0$ 后，可以构造出 $\boldsymbol{r}_1=A\boldsymbol{x}_1-\boldsymbol{b}=\boldsymbol{r}_0+\alpha_0A\boldsymbol{r}_0$，于是可以在直线 $\boldsymbol{x}=\boldsymbol{x}_1+\alpha\boldsymbol{r}_1$ 上求一点 $\boldsymbol{x}_2=\boldsymbol{x}_1+\alpha_1\boldsymbol{r}_1$ 使得

$$F_1(\boldsymbol{x}_1+\alpha_1\boldsymbol{r}_1)<F_1(\boldsymbol{x}_1+\alpha\boldsymbol{r}_1),$$

对任意 $\alpha\neq\alpha_1$ 的实数成立. 这样的

$$\alpha_1=-(\boldsymbol{r}_1, \boldsymbol{r}_1)/(A\boldsymbol{r}_1, \boldsymbol{r}_1),$$

依此类推，有计算程式:

给定 $\boldsymbol{x}_0$,

$$\begin{cases}\boldsymbol{r}_k=A\boldsymbol{x}_k-\boldsymbol{b},\\ \alpha_k=-(\boldsymbol{r}_k, \boldsymbol{r}_k)/(A\boldsymbol{r}_k, \boldsymbol{r}_k),\\ \boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\alpha_k\boldsymbol{r}_k, \quad k=0, 1, 2, \cdots,\end{cases}\tag{9}$$

从(9)构造出序列 $\{\boldsymbol{x}_k\}$ 的方法就称为斜量法，因为极小是在 $\boldsymbol{r}_k$ 的方向上取的，而

$$\boldsymbol{r}_k=\frac{1}{2}\,\text{grad}\,F_1(\boldsymbol{x})\Big|_{\boldsymbol{x}=\boldsymbol{x}_k},$$

因此称为斜量法.

斜量又称梯度(gradient)，它的几何意义是使 $F_1(\boldsymbol{x})$ 在某点邻近变化最快的方向，因此从 $F_1(\boldsymbol{x}_0+\alpha\boldsymbol{r}_0)$ 求极小，希望比其他方向 $\boldsymbol{p}$ 上求 $F_1(x_0+\alpha\boldsymbol{p})$ 极小，使得 $F_1(\boldsymbol{x})$ 下降得更快一点.

下面定理回答了从(9)导出的序列 $\{\boldsymbol{x}_k\}$ 的收敛性问题.

定理 1.2 设 A 的特征值为 λ_1、λ_2、…、λ_n,

$$\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n,$$

则

$$\|\boldsymbol{x}_k-\tilde{\boldsymbol{x}}\|\leqslant\sqrt{\frac{\lambda_n}{\lambda_1}}\left(\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}\right)^k\|\boldsymbol{x}_0-\tilde{\boldsymbol{x}}\|. \tag{10}$$

为了证明(10)式成立，先要证明两个引理. 设矩阵 A 的特征值 λ_i 对应的单位特征向量为 $\boldsymbol{y}_i$，并且 $\boldsymbol{y}_1$、$\boldsymbol{y}_2$、…、$\boldsymbol{y}_n$ 是一组标准正交向量组.

引理 1.1 在定理 1.2 的假设下，成立下列不等式

$$\lambda_1(\boldsymbol{x},\ \boldsymbol{x})\leqslant(A\boldsymbol{x},\ \boldsymbol{x})\leqslant\lambda_n(\boldsymbol{x},\ \boldsymbol{x}).$$

证明 令 $\boldsymbol{x}=\sum\limits_{j=1}^{n}\beta_j\boldsymbol{y}_j$，于是

$$(A\boldsymbol{x},\ \boldsymbol{x})=\sum_{j=1}^{n}\lambda_j\beta_j^2,$$

$$\lambda_1(\boldsymbol{x},\ \boldsymbol{x})=\sum_{j=1}^{n}\lambda_1\beta_j^2,$$

$$\lambda_n(\boldsymbol{x},\ \boldsymbol{x})=\sum_{j=1}^{n}\lambda_n\beta_j^2,$$

故 $\lambda_1(\boldsymbol{x},\ \boldsymbol{x})\leqslant(A\boldsymbol{x},\ \boldsymbol{x})\leqslant\lambda_n(\boldsymbol{x},\ \boldsymbol{x})$. 证毕.

引理 1.2 设 $\varphi(\lambda)$ 是一个 λ 的多项式，则

$$[\varphi(A)\boldsymbol{x},\ \varphi(A)\boldsymbol{x}]\leqslant\max_i\varphi(\lambda_i)^2[\boldsymbol{x},\ \boldsymbol{x}].$$

证明 令 $\boldsymbol{x}=\sum_{j=1}^{n}\beta_j\boldsymbol{y}_j$,

$$\varphi(A)\boldsymbol{x}=\sum_{j=1}^{n}\varphi(\lambda_j)\beta_j\boldsymbol{y}_j,$$

$$[\varphi(A)\boldsymbol{x},\ \varphi(A)\boldsymbol{x}]=\sum_{j=1}^{n}\lambda_j\varphi(\lambda_j)^2\beta_j^2$$

$$\leqslant\max_i\varphi(\lambda_i)^2\sum_{j=1}^{n}\lambda_j\beta_j^2=\max_i\varphi(\lambda_i)^2[\boldsymbol{x},\ \boldsymbol{x}].$$ 证毕.

现在来证明定理 1.2. 若 $\boldsymbol{r}_0=0$, (10)式自然成立，假设 $\boldsymbol{r}_0\neq 0$, 由 $F_1(\boldsymbol{x}_1)\leqslant F_1(\boldsymbol{x}_0+\alpha\boldsymbol{r}_0)$ 知

$$[\boldsymbol{x}_1-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_1-\tilde{\boldsymbol{x}}]\leqslant[\boldsymbol{x}_0+\alpha\boldsymbol{r}_0-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_0+\alpha\boldsymbol{r}_0-\tilde{\boldsymbol{x}}]$$
$$=[(I+\alpha A)(\boldsymbol{x}_0-\tilde{\boldsymbol{x}}),\ (I+\alpha A)(\boldsymbol{x}_0-\tilde{\boldsymbol{x}})],$$

对于多项式 $1+\alpha\lambda$, 应用引理 1.2, 得到

$$[\boldsymbol{x}_1-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_1-\tilde{\boldsymbol{x}}]\leqslant\max_i(1+\alpha\lambda_i)^2[\boldsymbol{x}_0-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_0-\tilde{\boldsymbol{x}}]$$
$$\leqslant\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}(1+\alpha\lambda)^2[\boldsymbol{x}_0-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_0-\tilde{\boldsymbol{x}}],$$

这一不等式对任何实数 α 都成立的，因此对特别选取的 $\alpha=-\frac{2}{\lambda_1+\lambda_n}$, 也成立不等式

$$[\boldsymbol{x}_1-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_1-\tilde{\boldsymbol{x}}]$$
$$\leqslant\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}\left(1-\frac{2\lambda}{\lambda_1+\lambda_n}\right)^2[\boldsymbol{x}_0-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_0-\tilde{\boldsymbol{x}}],$$

但
$$\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}\left(1-\frac{2\lambda}{\lambda_1+\lambda_n}\right)^2=\left(\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}\right)^2,$$

所以

$$[\boldsymbol{x}_1-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_1-\tilde{\boldsymbol{x}}]\leqslant\left(\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}\right)^2[\boldsymbol{x}_0-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_0-\tilde{\boldsymbol{x}}],$$

这样的关系，对于 $\boldsymbol{x}_k$、$\boldsymbol{x}_{k-1}$ 之间也可以同样证明成立的，即

$$[\boldsymbol{x}_k-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_k-\tilde{\boldsymbol{x}}]\leqslant\left(\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}\right)^2[\boldsymbol{x}_{k-1}-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_{k-1}-\tilde{\boldsymbol{x}}],$$

而 k 又是任意自然数，故

$$[\boldsymbol{x}_k-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_k-\tilde{\boldsymbol{x}}]\leqslant\left(\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}\right)^{2k}[\boldsymbol{x}_0-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_0-\tilde{\boldsymbol{x}}],$$

再利用引理1.1就可获得估计式(10). 证毕.

在证明中取 $\alpha=-\dfrac{2}{\lambda_1+\lambda_n}$，它是下列极大极小问题的解，

$$\min_{\alpha}\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|1+\alpha\lambda|$$

即对任何实数 α 有

$$\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}\left|1-\frac{2}{\lambda_1+\lambda_n}\lambda\right|\leqslant\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|1+\alpha\lambda|, \qquad (11)$$

等式只有在 $\alpha=-2/(\lambda_1+\lambda_n)$ 时达到，这可从图1知道. 由(11)式可知取 $\alpha=-2/(\lambda_1+\lambda_n)$ 可以得到比较小的误差估计，因此在证明中取这一个值.

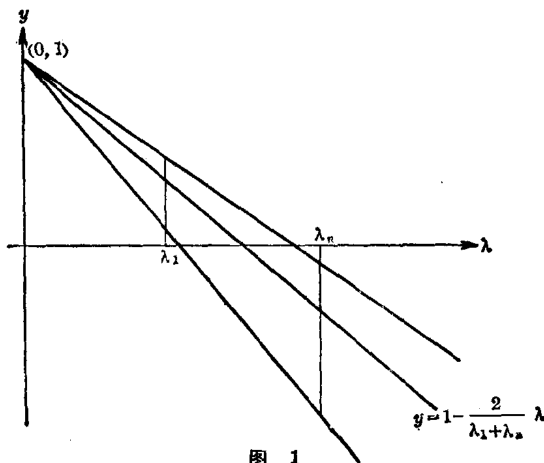

图 1

因为 $0<\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}<1$，从定理 1.2 知斜量法所得的序列 $\{x_k\}$ 是收敛的.

称 $p=\|A\|\|A^{-1}\|$ 为系数矩阵 A 的条件数. 当 A 是对称正定时 $p=\lambda_n/\lambda_1$，因此 $\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}=\frac{p-1}{p+1}$，从而可知当 p 很大时 $\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}$ 接近于 1. 说明矩阵 A 的条件数很大时(此时称矩阵 A 是病态的)，(10)式右端收敛于 0 是很慢的，实际计算的经验也是这样，此时 x_k 收敛于 $\tilde{x}$ 的速度也是很慢.

为了对系数矩阵 A 的条件数有个数量的概念，我们用差分方法求解正方形 $0\leqslant x, y\leqslant 1$ 上的 Poisson 方程第一类边值问题为例，若差分步长为 $h=1/m$，则系数矩阵为 $(m-1)^2$ 阶，它的 $(m-1)^2$ 个特征值为

$$\lambda_{p,q}=1-\frac{1}{2}(\cos p\pi h+\cos q\pi h)\qquad p、q=1, 2, \cdots, m-1,$$

故
$$\lambda_1=1-\cos h\pi=2\sin^2\frac{h\pi}{2},$$

$$\lambda_n=1-\cos(m-1)h\pi=2\sin^2\frac{(m-1)\pi}{2m},$$

于是
$$p=\frac{\lambda_n}{\lambda_1}\approx 4/(h\pi)^2,$$

$$h=1/10,\ p\approx 39.87,\ \frac{p-1}{p+1}=0.95,$$

$$h=1/50,\ p\approx 1000,\ \frac{p-1}{p+1}=0.998,$$

$$h=1/100,\ p\approx 4\times 10^3,\ \frac{p-1}{p+1}=0.9995,$$

当 $h=1/10$，$\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}=0.95$，如果要

$$\left(\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}\right)^k\leqslant 10^{-3},$$

$$k\geqslant -3/\log 0.95=134.65,$$

这告诉我们即使对于 $h=1/10$，对 81 阶的系数矩阵，要使

$$\|x_k-\tilde{x}\|\leqslant 10^{-3}\sqrt{\frac{\lambda_n}{\lambda_1}}\|x_0-\tilde{x}\|,$$

也要迭代 135 次；这就表示斜量法的收敛速度显得太慢了，因此在目前求解线性代数方程组时，不再采用斜量法. 不过有关斜量法的一些想法对于构造新的方法却仍然是有价值的.

§2 多步斜量法

上节所介绍的斜量法，它从 $\boldsymbol{x}_0$ 求 $\boldsymbol{x}_1$，是在直线 $\boldsymbol{x}_0+\alpha\boldsymbol{r}_0$ 上找的. 同样可以考虑，从 $\boldsymbol{x}_0$ 求 $\boldsymbol{x}_1$，在一个平面 $\boldsymbol{x}_0+\alpha\boldsymbol{p}+\beta\boldsymbol{q}$ 上找，也即 $\boldsymbol{x}_1=\boldsymbol{x}_0+\alpha_0\boldsymbol{p}+\beta_0\boldsymbol{q}$，并且不等式

$$F_1(\boldsymbol{x}_1)\leqslant F_1(\boldsymbol{x}_0+\alpha\boldsymbol{p}+\beta\boldsymbol{q})$$

对一切实数 α、β 成立. 利用定理 1.1 知道这样的 α_0、β_0 由下列方程组所确定：

$$\begin{cases}[\boldsymbol{x}_0+\alpha_0\boldsymbol{p}+\beta_0\boldsymbol{q}-\tilde{\boldsymbol{x}},\ \boldsymbol{p}]=0,\\ [\boldsymbol{x}_0+\alpha_0\boldsymbol{p}+\beta_0\boldsymbol{q}-\tilde{\boldsymbol{x}},\ \boldsymbol{q}]=0.\end{cases}$$

只要 $\boldsymbol{p}$，$\boldsymbol{q}$ 线性无关，上述方程唯一确定 α_0 和 β_0.

特别取 $\boldsymbol{p}=\boldsymbol{r}_0$，$\boldsymbol{q}$ 为另一与 $\boldsymbol{r}_0$ 线性无关的向量时，$\boldsymbol{x}_1=\boldsymbol{x}_0+\alpha_0\boldsymbol{r}_0+\beta_0\boldsymbol{q}$，有可能比斜量法中所确定的 $\boldsymbol{x}_1$ 要好，不可能比它差.

另外在§1 估计斜量法的收敛速度时，用到等式

$$\boldsymbol{x}_0+\alpha\boldsymbol{r}_0-\tilde{\boldsymbol{x}}=(1+\alpha A)(\boldsymbol{x}_0-\tilde{\boldsymbol{x}}),$$

从而把问题归结为考虑多项式 $1+\alpha\lambda$ 在区间 $[\lambda_1,\ \lambda_n]$ 上的极大极小问题

$$\min_{\alpha} \max_{\lambda_1 \leq \lambda \leq \lambda_n} |1+\alpha\lambda|.$$

如果我们仍要把在平面 $\boldsymbol{x}_0+\alpha\boldsymbol{p}+\beta\boldsymbol{q}$ 上找 $\boldsymbol{x}_1$ 的问题，归结为多项式的极大极小问题，自然会想到取 $\boldsymbol{q}=A\boldsymbol{r}_0$，这样

$$\boldsymbol{x}_0+\alpha\boldsymbol{p}+\beta\boldsymbol{q}=\boldsymbol{x}_0+\alpha_0\boldsymbol{r}_0+\beta A\boldsymbol{r}_0,$$

而

$$\begin{aligned}\boldsymbol{x}_0+\alpha\boldsymbol{r}_0+\beta A\boldsymbol{r}_0-\tilde{\boldsymbol{x}}&=\boldsymbol{x}_0-\tilde{\boldsymbol{x}}+\alpha\boldsymbol{r}_0+\beta A\boldsymbol{r}_0\\&=(I+\alpha A+\beta A^2)(\boldsymbol{x}_0-\tilde{\boldsymbol{x}}).\end{aligned}$$

如果 $\boldsymbol{r}_0$ 与 $A\boldsymbol{r}_0$ 线性无关，就可以唯一确定

$$\boldsymbol{x}_1=\boldsymbol{x}_0+\alpha_0\boldsymbol{r}_0+\beta_0A\boldsymbol{r}_0,$$

如果 $\boldsymbol{r}_0$ 与 $A\boldsymbol{r}_0$ 线性相关，我们来分析一下，会发生什么？假定 $\boldsymbol{r}_0\neq 0$，此时必有 $a\neq 0$，而

$$A\boldsymbol{r}_0=a\boldsymbol{r}_0,$$

因为 A^{-1} 存在，故

$$\boldsymbol{r}_0=a(\boldsymbol{x}_0-\tilde{\boldsymbol{x}}),$$

或

$$\tilde{\boldsymbol{x}}=\boldsymbol{x}_0-\frac{1}{a}\boldsymbol{r}_0,$$

说明解向量 $\tilde{\boldsymbol{x}}$ 在直线 $\boldsymbol{x}_0+\alpha\boldsymbol{r}_0$ 上，因此 $F_1(\boldsymbol{x}_0+\alpha\boldsymbol{r}_0+\beta A\boldsymbol{r}_0)$ 上的极小一定在 $\tilde{\boldsymbol{x}}=\boldsymbol{x}_0-\frac{1}{a}\boldsymbol{r}_0$ 上达到． $\boldsymbol{x}_1$ 即为解 $\tilde{\boldsymbol{x}}$． 不管 $\boldsymbol{r}_0$ 与 $A\boldsymbol{r}_0$ 是否线性无关，α_0、β_0 都可由

$$\begin{cases}[\boldsymbol{x}_0+\alpha_0\boldsymbol{r}_0+\beta_0A\boldsymbol{r}_0-\tilde{\boldsymbol{x}},\ \boldsymbol{r}_0]=0,\\ [\boldsymbol{x}_0+\alpha_0\boldsymbol{r}_0+\beta_0A\boldsymbol{r}_0-\tilde{\boldsymbol{x}},\ A\boldsymbol{r}_0]=0\end{cases}$$

确定．一般的计算程式为：

取定一个 $\boldsymbol{x}_0$，计算 $\boldsymbol{r}_0=A\boldsymbol{x}_0-\boldsymbol{b}$，计算 $A\boldsymbol{r}_0$，求解 α_k、β_k：

$$\begin{cases}[\boldsymbol{r}_k,\ \boldsymbol{r}_k]\alpha_k+[A\boldsymbol{r}_k,\ \boldsymbol{r}_k]\beta_k=-[\boldsymbol{x}_k-\tilde{\boldsymbol{x}},\ \boldsymbol{r}_k],\\ [\boldsymbol{r}_k,\ A\boldsymbol{r}_k]\alpha_k+[A\boldsymbol{r}_k,\ A\boldsymbol{r}_k]\beta_k=-[\boldsymbol{x}_k-\tilde{\boldsymbol{x}},\ A\boldsymbol{r}_k],\\ \boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\alpha_k\boldsymbol{r}_k+\beta_kA\boldsymbol{r}_k,\\ \boldsymbol{r}_{k+1}=A\boldsymbol{x}_{k+1}-\boldsymbol{b},\quad k=0,\ 1,\ 2,\ \cdots,\end{cases}\tag{12}$$

这样确定 $\{\boldsymbol{x}_k\}$ 的方法，称为二步斜量法．注意(12)中方程的右端是 $-(\boldsymbol{r}_k,\ \boldsymbol{r}_k)$、$-(\boldsymbol{r}_k,\ A\boldsymbol{r}_k)$，因此是可以计算的已知量．

记 $$E_2=\min_{\alpha,\beta}\ \max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|1+\alpha\lambda+\beta\lambda^2|,$$

跟定理 1.2 的证明办法一样，可以证明二步斜量法有收敛速度的估计

$$\|\boldsymbol{x}_k-\tilde{\boldsymbol{x}}\|\leqslant\sqrt{\frac{\lambda_n}{\lambda_1}}\,E_2^k\|\boldsymbol{x}_0-\tilde{\boldsymbol{x}}\|.$$

按照二步斜量法的思路，对于取定的 $\boldsymbol{x}_0$，构造 $\boldsymbol{r}_0$、$A\boldsymbol{r}_0$、…，$A^{l-1}\boldsymbol{r}_0$，然后在 l 维超平面

$$\boldsymbol{x}_0+\alpha^{(1)}\boldsymbol{r}_0+\alpha^{(2)}A\boldsymbol{r}_0+\cdots+\alpha^{(l)}A^{l-1}\boldsymbol{r}_0$$

上(这里 $\alpha^{(1)}$、$\alpha^{(2)}$、…、$\alpha^{(l)}$ 是任意实数)找一个点 $\boldsymbol{x}_1$，

$$\boldsymbol{x}_1=\boldsymbol{x}_0+\alpha_0^{(1)}\boldsymbol{r}_0+\alpha_0^{(2)}A\boldsymbol{r}_0+\cdots+\alpha_0^{(l)}A^{l-1}\boldsymbol{r}_0,$$

使得

$$F_1(\boldsymbol{x}_1)\leqslant F_1(\boldsymbol{x}_0+\alpha^{(1)}\boldsymbol{r}_0+\cdots+\alpha^{(l)}A^{l-1}\boldsymbol{r}_0)$$

对所有实数 $\alpha^{(1)}$、$\alpha^{(2)}$、…、$\alpha^{(l)}$ 成立．

由定理 1.1 知道这样的 $\alpha_0^{(1)}$、$\alpha_0^{(2)}$、…、$\alpha_0^{(l)}$ 由下列方程组确定

$$\begin{cases}[\boldsymbol{x}_0+\alpha_0^{(1)}\boldsymbol{r}_0+\alpha_0^{(2)}A\boldsymbol{r}_0+\cdots+\alpha_0^{(l)}A^{l-1}\boldsymbol{r}_0-\tilde{\boldsymbol{x}},\ \boldsymbol{r}_0]=0,\\ [\boldsymbol{x}_0+\alpha_0^{(1)}\boldsymbol{r}_0+\alpha_0^{(2)}A\boldsymbol{r}_0+\cdots+\alpha_0^{(l)}A^{l-1}\boldsymbol{r}_0-\tilde{\boldsymbol{x}},\ A\boldsymbol{r}_0]=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ [\boldsymbol{x}_0+\alpha_0^{(1)}\boldsymbol{r}_0+\alpha_0^{(2)}A\boldsymbol{r}_0+\cdots+\alpha_0^{(0)}A^{l-1}\boldsymbol{r}_0-\tilde{\boldsymbol{x}},\ A^{l-1}\boldsymbol{r}_0]=0,\end{cases}$$

当 $\boldsymbol{r}_0$, $A\boldsymbol{r}_0$, …, $A^{l-1}\boldsymbol{r}_0$ 线性无关时，上面的方程组唯一确定一组 $\alpha_0^{(1)}$、$\alpha_0^{(2)}$、…、$\alpha_0^{(l)}$；当 $\boldsymbol{r}_0$、$A\boldsymbol{r}_0$、…、$A^{l-1}\boldsymbol{r}_0$ 线性相关时，$\alpha_0^{(1)}$、$\alpha_0^{(2)}$、…、$\alpha_0^{(l)}$ 不唯一，但每组解都对应 $\boldsymbol{x}_1=\tilde{\boldsymbol{x}}$.

由此得到 l 步斜量法的计算程式：

给定 $\boldsymbol{x}_0$，计算 $\boldsymbol{r}_0=A\boldsymbol{x}_0-\boldsymbol{b}$, $A\boldsymbol{r}_0$, …, $A^{l-1}\boldsymbol{r}_0$, 求解 $\alpha_k^{(1)}$,

$\alpha_k^{(2)}, \cdots, \alpha_k^{(l)}$:

$$\begin{cases}[\boldsymbol{r}_k, \boldsymbol{r}_k]\alpha_k^{(1)}+[A\boldsymbol{r}_k, \boldsymbol{r}_k]\alpha_k^{(2)}+\cdots+[A^{l-1}\boldsymbol{r}_k, \boldsymbol{r}_k]\alpha_k^{(l)} \\ \quad =-[\boldsymbol{x}_k-\widetilde{\boldsymbol{x}}, \boldsymbol{r}_k], \\ [\boldsymbol{r}_k, A\boldsymbol{r}_k]\alpha_k^{(1)}+[A\boldsymbol{r}_k, A\boldsymbol{r}_k]\alpha_k^{(2)}+\cdots+[A^{l-1}\boldsymbol{r}_k, A\boldsymbol{r}_k]\alpha_k^{(l)} \\ \quad =-[\boldsymbol{x}_k-\widetilde{\boldsymbol{x}}, A\boldsymbol{r}_k], \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ [\boldsymbol{r}_k, A^{l-1}\boldsymbol{r}_k]\alpha_k^{(1)}+[A\boldsymbol{r}_k, A^{l-1}\boldsymbol{r}_k]\alpha_k^{(2)}+\cdots \\ \quad +[A^{l-1}\boldsymbol{r}_k, A^{l-1}\boldsymbol{r}_k]\alpha_k^{(l)}=-[\boldsymbol{x}_k-\widetilde{\boldsymbol{x}}, A^{l-1}\boldsymbol{r}_k], \\ \boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\alpha_k^{(1)}\boldsymbol{r}_k+\alpha_k^{(2)}A\boldsymbol{r}_k+\cdots+\alpha_k^{(l)}A^{l-1}\boldsymbol{r}_k, \\ \boldsymbol{r}_{k+1}=A\boldsymbol{x}_{k+1}-\boldsymbol{b}, \quad k=0, 1, 2, \cdots.\end{cases} \tag{13}$$

记

$$E_l=\min_{\alpha^{(1)},\alpha^{(2)},\cdots,\alpha^{(l)}}\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|1+\alpha^{(1)}\lambda+\alpha^{(2)}\lambda+\cdots+\alpha^{(l)}\lambda^l|,$$

同样对 l 步斜量法有收敛速度的估计

$$|\boldsymbol{x}_1-\widetilde{\boldsymbol{x}}|\leqslant E_l|\boldsymbol{x}_0-\widetilde{\boldsymbol{x}}|,$$

$$\|\boldsymbol{x}_k-\widetilde{\boldsymbol{x}}\|\leqslant\sqrt{\frac{\lambda_n}{\lambda_1}}E_l^k\|\boldsymbol{x}_0-\widetilde{\boldsymbol{x}}\|. \tag{14}$$

在§1已经知道

$$E_1=\min_{\alpha}\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|1+\alpha\lambda|=\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1},$$

对于 $l>1$, E_l 究竟是怎样一个量？这是一个需要进一步回答的问题. E_l 是从型如 $1+\alpha^{(1)}\lambda+\alpha^{(2)}\lambda^2+\cdots+\alpha^{(l)}\lambda^l$ 的多项式全体中考虑极大极小得到的量，也即是从所有 l 次多项式 $q_l(\lambda)$ 中，满足 $q_l(0)=1$ 的那些 l 次多项式中考虑极大极小得到的量.

下面先介绍一下契贝谢夫(П. Л. Чебышев)多项式.

$$T_l(x)=\frac{(x+\sqrt{x^2-1})^l+(x-\sqrt{x^2-1})^l}{2}$$

是一个首项系数为 2^{l-1} 的 l 次多项式，称为 l 次契贝谢夫多

项式. 如果把自变量 x 的范围限制在 $[-1, 1]$ 上, 那么可以作变换 $x=\cos\theta$, $\theta\in[0, \pi]$, θ 与 x 之间建立起一一对应.

$$T_l(\cos\theta)=\frac{(\cos\theta+\sqrt{\cos^2\theta-1})^l+(\cos\theta-\sqrt{\cos^2\theta-1})^l}{2}$$

$$=\frac{(\cos\theta+i\sin\theta)^l+(\cos\theta-i\sin\theta)^l}{2}=\cos l\theta,$$

因此当 $x\in[-1, 1]$ 时,

$$T_l(x)=\cos l\arccos x,$$

从而也有 $|T_l(x)|\leqslant 1$.

取 $x_k=\cos\dfrac{k\pi}{l}\in[-1, 1]$, 有

$$T_l(x_k)=\cos l\arccos\frac{k\pi}{l}=\cos k\pi=(-1)^k,$$

对于 $k=0、1、2、\cdots、l$, $x_0、x_1、\cdots、x_l$ 是 $l+1$ 个从右到左排列在区间 $[-1, 1]$ 上不同的点, 它们轮流使 $T_l(x)$ 取正 1 或负 1, 而 1 是 $T_l(x)$ 在 $[-1, 1]$ 上的最大值, 这样的一组点称为正负交替偏差点. 有 $l+1$ 个正负交替偏差点是 l 次契贝谢夫多项式的一个重要性质.

今考虑如下 l 次多项式

$$P_l(\lambda)=T_l\left(\frac{2\lambda}{\lambda_n-\lambda_1}-\frac{\lambda_1+\lambda_n}{\lambda_n-\lambda_1}\right)\Big/T_l\left(-\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right),$$

显然有 $P_l(0)=1$, 并且当 $\lambda\in[\lambda_1, \lambda_n]$ 时,

$$x=\frac{2\lambda}{\lambda_n-\lambda_1}-\frac{\lambda_1+\lambda_n}{\lambda_n-\lambda_1}\in[-1, 1],$$

因此当 $\lambda\in[\lambda_1, \lambda_n]$ 时

$$|P_l(\lambda)|\leqslant 1\Big/\left|T_l\left(-\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)\right|,$$

将 $[-1, 1]$ 上的 x_k 与 $[\lambda_1, \lambda_n]$ 中的 η_k 建立起对应, 即

$$x_k=\frac{2\eta_k}{\lambda_n-\lambda_1}-\frac{\lambda_1+\lambda_n}{\lambda_n-\lambda_1},$$

或
$$\eta_k=\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}+\cos\frac{k\pi}{l}\right)\frac{\lambda_n-\lambda_1}{2},$$

有 $P_l(\eta_k)=(-1)^k\Big/T_l\left(-\frac{\lambda_1+\lambda_n}{\lambda_n-\lambda_1}\right)$，这说明 $\eta_k(k=0,1,\cdots,l)$是 $P_l(t)$ 在$[\lambda_1, \lambda_n]$中的 $l+1$ 个正负交替的偏差点.

定理1.3 $E_l=\max\limits_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|P_l(\lambda)|=1\Big/T_l\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)$.

证明 由 E_l 的定义可知:

$$E_l\leqslant\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|P_l(\lambda)|,$$

如果 $E_l<\max\limits_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|P_l(\lambda)|$，那么必有一个 l 次多项式 $Q(\lambda)$，且 $Q(0)=1$，使得

$$\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|Q(\lambda)|<\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|P_l(\lambda)|, \tag{15}$$

于是 $P_l(\lambda)-Q(\lambda)$ 在 $\eta_k(k=0, 1, \cdots, l)$ 的符号与 $(-1)^k\Big/T_k\left(-\frac{\lambda_1+\lambda_n}{\lambda_n-\lambda_1}\right)$相同，从而 $P_l(\lambda)-Q(\lambda)$在$[\lambda_1, \lambda_n]$中至少有 l 个零点，另外 0 不在区间$[\lambda_1,\lambda_n]$中，而

$$P_l(0)-Q(0)=1-1=0,$$

因此 0 也是 $P_l(\lambda)-Q(\lambda)$的一个零点，这样 $P_l(\lambda)-Q(\lambda)$ 至少有 $l+1$ 个零点，因为这是一个 l 次多项式，故必须

$$P_l(\lambda)-Q(\lambda)\equiv0,$$

此与(15)矛盾，故知

$$E_l=\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|P_l(\lambda)|,$$

但
$$\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|P_l(\lambda)|=1\Big/\left|T_l\left(-\frac{\lambda_1+\lambda_n}{\lambda_n-\lambda_1}\right)\right|,$$

而对于契贝谢夫多项式有

$$T_l\left(-\frac{\lambda_1+\lambda_n}{\lambda_n-\lambda_1}\right)=(-1)^l T_l\left(\frac{\lambda_1+\lambda_n}{\lambda_n-\lambda_1}\right),$$

故 $$E_l=1\Big/T_l\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right).$$ 证毕.

按 $T_l(x)$ 的定义

$$\begin{aligned}E_l&=1\Big/T_l\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)\\&=2\Big/\left[\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}+\sqrt{\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)^2-1}\right)^l\right.\\&\qquad\left.+\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}-\sqrt{\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)^2-1}\right)^l\right],\end{aligned}$$

当 $l=1$ 时 $$E_1=\frac{2}{\dfrac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}+\dfrac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}}=\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1},$$

与 §1 的结果相符.

定理1.3没有说明在 $1+\alpha^{(1)}\lambda+\cdots+\alpha^{(l)}\lambda^l$ 型的多项式中，是否只有一个 $P_l(\lambda)$，使得

$$\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}\left|1+\alpha^{(1)}\lambda+\alpha^{(2)}\lambda^2+\cdots+\alpha^{(l)}\lambda^l\right|=E_l.$$

这个问题的回答是肯定的.

定理1.4 如果 l 次多项式 $Q(\lambda)$，$Q(0)=1$，满足如下条件

$$\max_{\lambda_1\leqslant\lambda\leqslant\lambda_n}|Q(\lambda)|=E_l,$$

则 $Q(\lambda)\equiv P_l(\lambda)$.

证明 由定理1.3知 $P_l(\lambda)-Q(\lambda)$ 在 η_k 上或者同号于 $(-1)^k\Big/T_l\left(-\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)$，或者为0. 对于 $\eta_k(k=1,2,\cdots,l-1)$ 有 $\eta_k\in(\lambda_1,\lambda_n)$，因此当在这些内部的 η_k 如果

$$P_l(\eta_k)-Q(\eta_k)=0,$$

必有 $P_l'(\eta_k)-Q'(\eta_k)=0$，这是因为 η_k 是 $P_l(\lambda)$ 的极大(小)值，也是 $Q(\lambda)$ 的极大(小)值. 这时，η_k 是 $P_l(\lambda)-Q(\lambda)$ 的重根. 由此可知 $P_l(\lambda)-Q(\lambda)$ 在 $[\lambda_1,\lambda_n]$ 中至少有 l 个零点. 另外再由 $P_l(0)-Q(0)=0$，知 0 也是一个零点，从而可知 $P_l(\lambda)-Q(\lambda)$ 至少有 $l+1$ 个零点，故

$$Q(\lambda)\equiv P_l(\lambda). \quad \text{证毕.}$$

现在来比较斜量法与多步斜量法的收敛速度. 在计算中，主要的计算量是化在矩阵与向量的乘法上，我们以矩阵与向量乘法一次作为一个单位.

对斜量法来说从 $\boldsymbol{x}_0$ 求 $\boldsymbol{x}_1$ 要计算二次矩阵和向量的乘法，即 $\boldsymbol{r}_0=A\boldsymbol{x}_0-\boldsymbol{b}$，$A\boldsymbol{r}_0$，但对以后从 $\boldsymbol{x}_k$ 求 $\boldsymbol{x}_{k+1}$，每步只要计算一次矩阵和向量的乘法 $A\boldsymbol{r}_k$.

对于二步斜量法来说，从 $\boldsymbol{x}_0$ 求 $\boldsymbol{x}_1$ 要计算三次矩阵和向量的乘法，$\boldsymbol{r}_0=A\boldsymbol{x}_0-\boldsymbol{b}$、$A\boldsymbol{r}_0$、$A^2\boldsymbol{r}_0$，但对以后从 $\boldsymbol{x}_k$ 求 $\boldsymbol{x}_{k+1}$ 每步只要二次矩阵和向量的乘法 $A\boldsymbol{r}_k$、$A^2\boldsymbol{r}_k$.

如果我们称斜量法为一步斜量法，那么不算从 $\boldsymbol{x}_0$ 求 x_1，一般地从 $\boldsymbol{x}_k$ 求 $\boldsymbol{x}_{k+1}$，l 步斜量法需要计 l 次矩阵和向量的乘法. 因此 l 步斜量迭代一次，计算量相当于斜量法迭代 l 次. 因此我们要比较 E_l 与 E_1^l 究竟那个小，才能知道那个收敛速度快. 同样各种多步斜量法之间也要作比较，l 步斜量法与 k 步斜量法究竟那个快，就要比较 $(E_l)^{1/l}$ 与 $(E_k)^{1/k}$ 究竟那个小. 下面定理回答了这个问题.

定理1.5 设 $T_s(\sigma)=\dfrac{(\sigma+\sqrt{\sigma^2-1})^s+(\sigma-\sqrt{\sigma^2-1})^s}{2}$ 是 s 次契贝谢夫多项式. 若 $\sigma>1$，$k>l$，则

$$(T_k(\sigma))^{1/k}>(T_l(\sigma))^{1/l}.$$

证明 令 $\alpha=\sigma+\sqrt{\sigma^2-1}$，$\beta=\sigma-\sqrt{\sigma^2-1}$，于是

$$\theta=\beta/\alpha<1,$$

而 $T_s(\sigma)=\alpha^s\dfrac{1+\theta^s}{2}$, $(T_s(\sigma))^{1/s}=\alpha\left(\dfrac{1+\theta^s}{2}\right)^{1/s}$, 我们下面证明 $(T_s(\sigma))^{1/s}$ 是 s 的单调上升函数，为此只要证明

$$f(s)=\left(\frac{1+\theta^s}{2}\right)^{1/s}$$

是 s 的单调上升函数.

实际上

$$\begin{aligned}
f'(s)&=f(s)\frac{d}{ds}\left[\frac{1}{s}\ln\frac{1+\theta^s}{2}\right]\\
&=f(s)\left\{-\frac{1}{s^2}\ln\frac{1+\theta^s}{2}+\frac{1}{s}\,\frac{2}{1+\theta^s}\times\frac{1}{2}\theta^s\ln\theta\right\}\\
&=f(s)\left\{-\frac{1}{s^2}\ln\frac{1+\theta^s}{2}+\frac{s\theta^s\ln\theta}{s^2(1+\theta^s)}\right\}\\
&=f(s)\frac{(1+\theta^s)\ln 2-(1+\theta^s)\ln(1+\theta^s)+\theta^s\ln\theta^s}{s^2(1+\theta^s)}.
\end{aligned}$$

为了判明 $f'(s)$ 的符号，我们只要研究上述等式右边的分子符号，如果记 $\eta=\theta^s$，于是知 $\eta\in[0,1)$，上述分子为

$$\begin{aligned}
g(\eta)&=(1+\eta)\ln 2-(1+\eta)\ln(1+\eta)+\eta\ln\eta,\\
g'(\eta)&=\ln 2-\ln(1+\eta)-1+\ln\eta+1\\
&=\ln\frac{2\eta}{1+\eta}<0,\quad \text{当 }\eta\in(0,1),\\
g'(1)&=0,
\end{aligned}$$

再由 $g(0)=\ln 2$, $g(1)=0$, 利用微分学中值定理就知 $g(\eta)>0$ 当 $\eta\in[0,1)$. 再利用 $f(s)>0$, $s^2(1+\theta^s)>0$ 得到

$$f'(s)>0. \text{ 证毕.}$$

推论 当 $k>l$, $(E_k)^{\frac{1}{k}}<(E_l)^{\frac{1}{l}}$.

证明 $\sigma=\dfrac{\lambda_1+\lambda_n}{\lambda_n-\lambda_1}>1$,

而 $E_s=1/T_s(\sigma)$，利用定理 1.5 即可证明推论.

从定理1.5知道，多步斜量法比斜量法要快，而且步数越高越快，但是步数越高(13)中的方程组阶数也越高，需要更多的存贮单元，需要更多的计算量求解也越麻烦，下节介绍的共轭斜量法，避免了解(13)中的那样方程组。

§3 共轭斜量法

给定向量 $\boldsymbol{x}_0$, $\boldsymbol{r}_0=A\boldsymbol{x}_0-\boldsymbol{b}$，可以得到序列

$$\boldsymbol{r}_0、A\boldsymbol{r}_0、\cdots、A^l\boldsymbol{r}_0,\quad l<n,$$

假定这组向量是线性无关的，记子空间 $m_l=\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \cdots, A^l\boldsymbol{r}_0\}$，将这组向量按内积 $[\boldsymbol{x}, \boldsymbol{y}]$ 正交化，这样的正交化也称为 A-正交化，于是得到一组 A-正交化的序列

$$\boldsymbol{q}_0、\boldsymbol{q}_1、\cdots、\boldsymbol{q}_l,$$

假定正交化过程是如下那样次序进行的：

$$\begin{cases}\boldsymbol{q}_0=\boldsymbol{r}_0,\ \boldsymbol{q}_1=A\boldsymbol{r}_0+\sigma_0^{(0)}\boldsymbol{q}_0,\\ \boldsymbol{q}_2=A^2\boldsymbol{r}_0+\sigma_0^{(1)}\boldsymbol{q}_0+\sigma_1^{(1)}\boldsymbol{q}_1,\\ \text{一般}\\ \boldsymbol{q}_i=A^i\boldsymbol{r}_0+\sigma_0^{(i-1)}\boldsymbol{q}_0+\sigma_1^{(i-1)}\boldsymbol{q}_1+\cdots+\sigma_{i-1}^{(i-1)}\boldsymbol{q}_{i-1},\end{cases}\tag{16}$$

易知 $\boldsymbol{q}_i=0$ 的充要条件是 $\boldsymbol{r}_0、A\boldsymbol{r}_0、\cdots、A^i\boldsymbol{r}_0$ 线性相关. 由 A-正交化的要求，有

$$\begin{aligned}\sigma_j^{(i-1)}&=-[A^i\boldsymbol{r}_0, \boldsymbol{q}_j]/[\boldsymbol{q}_j, \boldsymbol{q}_j]\\&=-(A\boldsymbol{q}_j, A^i\boldsymbol{r}_0)/(A\boldsymbol{q}_j, \boldsymbol{q}_j),\end{aligned}$$

再由 $[\boldsymbol{q}_i, \boldsymbol{q}_j]=0$，知负的 $\sigma_0^{(i-1)}\boldsymbol{q}_0+\cdots+\sigma_{i-1}^{(i-1)}\boldsymbol{q}_{i-1}$ 是 $A^i\boldsymbol{r}_0$，在子空间 $\{\boldsymbol{q}_0, \boldsymbol{q}_1, \cdots, \boldsymbol{q}_{i-1}\}$ 上的最佳逼近.

从正交化的过程可知

$$\{\boldsymbol{q}_0, \boldsymbol{q}_1, \cdots, \boldsymbol{q}_l\}\equiv\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \cdots, A^l\boldsymbol{r}_0\}=m_l.$$

对于向量 $\boldsymbol{x}_0-\tilde{\boldsymbol{x}}$ 在子空间 m_l 上找一个向量 $\boldsymbol{y}$，使得

$[\boldsymbol{x}_0-\tilde{\boldsymbol{x}}+\boldsymbol{y},\ \boldsymbol{x}_0-\tilde{\boldsymbol{x}}+\boldsymbol{y}]$达到极小. 利用定理1.1知道,这样的$\boldsymbol{y}$必须且只须满足方程

$$[\boldsymbol{x}_0-\tilde{\boldsymbol{x}}+\boldsymbol{y},\ \boldsymbol{q}_i]=0,\quad i=0,\ 1,\ 2,\ \cdots,l. \tag{17}$$

若记

$$\boldsymbol{y}=\sum_{i=0}^{l}\alpha_i\boldsymbol{q}_i,$$

由(17)和$\boldsymbol{q}_0,\ \boldsymbol{q}_1,\ \cdots,\ \boldsymbol{q}_l$的正交性,得到

$$\alpha_i=-[\boldsymbol{x}_0-\tilde{\boldsymbol{x}},\ \boldsymbol{q}_i]/[\boldsymbol{q}_i,\ \boldsymbol{q}_i]=-(\boldsymbol{r}_0,\ \boldsymbol{q}_i)/(A\boldsymbol{q}_i,\ \boldsymbol{q}_i).$$

因为

$$[\boldsymbol{x}_0-\tilde{\boldsymbol{x}}+\boldsymbol{y},\ \boldsymbol{x}_0-\tilde{\boldsymbol{x}}+\boldsymbol{y}]=F_1(\boldsymbol{x}_0+\boldsymbol{y}),$$

故知这样得到的$\boldsymbol{y}$使得$\boldsymbol{x}_0+\boldsymbol{y}$即为$l+1$步斜量法中得到的$\boldsymbol{x}_1$,从而有

$$\left|\tilde{\boldsymbol{x}}-\boldsymbol{x}_0-\sum_{i=0}^{l}\alpha_i\boldsymbol{q}_i\right|\leqslant E_{l+1}|\tilde{\boldsymbol{x}}-\boldsymbol{x}_0|.$$

因为l可以是任意非负整数,注意到α_i的计算公式不随l变化而变化,因此我们可以得到如下的计算程式:

$$\begin{aligned}
&\boldsymbol{x}_1=\boldsymbol{x}_0+\alpha_0\boldsymbol{q}_0,\\
&\quad\alpha_0=-(\boldsymbol{r}_0,\ \boldsymbol{q}_0)/(A\boldsymbol{q}_0,\ \boldsymbol{q}_0),\\
&\boldsymbol{x}_2=\boldsymbol{x}_0+\alpha_0\boldsymbol{q}_0+\alpha_1\boldsymbol{q}_1=\boldsymbol{x}_1+\alpha_1\boldsymbol{q}_1,\\
&\quad\alpha_1=-(\boldsymbol{r}_0,\ \boldsymbol{q}_1)/(A\boldsymbol{q}_1,\ \boldsymbol{q}_1),\\
&\boldsymbol{x}_3=\boldsymbol{x}_0+\alpha_0\boldsymbol{q}_0+\alpha_1\boldsymbol{q}_1+\alpha_2\boldsymbol{q}_2=\boldsymbol{x}_2+\alpha_2\boldsymbol{q}_2,\\
&\quad\alpha_2=-(\boldsymbol{r}_0,\ \boldsymbol{q}_2)/(A\boldsymbol{q}_2,\ \boldsymbol{q}_2),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&\boldsymbol{x}_k=\boldsymbol{x}_{k-1}+\alpha_{k-1}\boldsymbol{q}_{k-1},\\
&\quad\alpha_{k-1}=-(\boldsymbol{r}_0,\ \boldsymbol{q}_{k-1})/(A\boldsymbol{q}_{k-1},\ \boldsymbol{q}_{k-1}).
\end{aligned}$$

这一计算$\{\boldsymbol{x}_k\}$的方法,就是共轭斜量法,因为用到了$\boldsymbol{r}_0$、$A\boldsymbol{r}_0$、$\cdots$、$A^l\boldsymbol{r}_0$的A-正交化组$\boldsymbol{q}_0$、$\boldsymbol{q}_1$、$\cdots$、$\boldsymbol{q}_l$,而两个向量A-正交又称为共轭,所以称为共轭斜量法. 并且得到估计式:

$$\|\boldsymbol{x}_k-\tilde{\boldsymbol{x}}\|\leqslant\sqrt{\frac{\lambda_n}{\lambda_1}}E_k\|\boldsymbol{x}_0-\tilde{\boldsymbol{x}}\|.$$

由正交化的过程中和计算 x_k 的过程中可以看到如果 $\boldsymbol{q}_{k-1}=0$, 那么计算就要中断, 但是恰巧在这时, $\boldsymbol{x}_{k-1}=\tilde{\boldsymbol{x}}$, 意味着计算已经完成. 下面来证明这一点, 首先由 $\boldsymbol{q}_{k-1}=0$, 知道 $\boldsymbol{r}_0$、$A\boldsymbol{r}_0$、…、$A^{k-1}\boldsymbol{r}_0$ 是线性相关的, 因此 $\boldsymbol{x}_0-\tilde{\boldsymbol{x}}$ 落在空间 $\{\boldsymbol{r}_0, A\boldsymbol{r}_0, \cdots, A^{k-2}\boldsymbol{r}_0\}$ 中, 也即 $\boldsymbol{x}_0-\tilde{\boldsymbol{x}}$ 落在空间 $\{\boldsymbol{q}_0, \boldsymbol{q}_1, \cdots, \boldsymbol{q}_{k-2}\}$ 中, 实际上从 $\boldsymbol{r}_0$、$A\boldsymbol{r}_0$、…、$A^{k-1}\boldsymbol{r}_0$ 线性相关, 知存在常数 ω_l, 使 $\sum\limits_{l=0}^{k-1}\omega_l A^l\boldsymbol{r}_0=0$, 等式两边作用 A^{-1}, 即得

$$\sum_{l=0}^{k-1}\omega_l A^l(\boldsymbol{x}_0-\tilde{\boldsymbol{x}})=0.$$

如果 $\omega_0\neq0$, 就知 $x_0-\tilde{x}$ 可用 r_0、Ar_0、…、$A^{k-2}r_0$ 来表示. 如果 $\omega_0=0$, $\omega_1\neq0$, 可以再作用 A^{-1}, 就得 $\boldsymbol{x}_0-\tilde{\boldsymbol{x}}$ 可用 $\boldsymbol{r}_0$、$A\boldsymbol{r}_0$、…、$A^{k-3}\boldsymbol{r}_0$ 线性组合表示. 以此类推, 可知这一结论成立. 这样就知

$$\tilde{\boldsymbol{x}}=\boldsymbol{x}_0+\beta_0\boldsymbol{q}_0+\beta_1\boldsymbol{q}_1+\cdots+\beta_{k-2}\boldsymbol{q}_{k-2},$$

但
$$\boldsymbol{x}_{k-1}=\boldsymbol{x}_0+\alpha_0\boldsymbol{q}_0+\alpha_1\boldsymbol{q}_1+\cdots+\alpha_{k-2}\boldsymbol{q}_{k-2},$$

而且
$$0\leqslant\left[\boldsymbol{x}_0+\sum_{i=0}^{k-2}\alpha_i\boldsymbol{q}_i-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_0+\sum_{i=0}^{k-2}\alpha_i\boldsymbol{q}_i-\tilde{\boldsymbol{x}}\right]$$
$$\leqslant\left[\boldsymbol{x}_0+\sum_{i=0}^{k-2}\beta_i\boldsymbol{q}_i-\tilde{\boldsymbol{x}},\ \boldsymbol{x}_0+\sum_{i=0}^{k-2}\beta_i\boldsymbol{q}_i-\tilde{\boldsymbol{x}}\right]=0,$$

故必须
$$\boldsymbol{x}_{k-1}=\boldsymbol{x}_0+\sum_{i=0}^{k-2}\alpha_i\boldsymbol{q}_i=\boldsymbol{x}_0+\sum_{i=0}^{k-2}\beta_i\boldsymbol{q}_i=\tilde{\boldsymbol{x}}.$$

这里还要指出一点: 从 (16) 那样的方法获得 A-正交化组, 无论从计算量角度还是从存贮量角度来说都是花费太大, 因为在计算 $\boldsymbol{q}_i$ 时, 要用到 $\boldsymbol{q}_0$、$\boldsymbol{q}_1$、…、$\boldsymbol{q}_{i-1}$, 后面这 i 个向量都要存放在快速存贮器中, 而且不能指望这些向量是稀疏的, 因此得有 $n\times i$ 个存贮单元, 当 i 较大时, 例如 $i=\frac{n}{2}$, 那么就需 $\frac{1}{2}n^2$ 个存贮单元, 这在实际计算中是太大的花费. 另外因为

用到 $\boldsymbol{q}_0$、$\boldsymbol{q}_1$、…、$\boldsymbol{q}_{i-1}$，相应的要计算 $\sigma_0^{(i-1)}$、$\sigma_1^{(i-1)}$、…、$\sigma_{i-1}^{(i-1)}$，也要付出相应的计算量．

因此要设法，在计算 $\boldsymbol{q}_i$ 时只用到少数几个向量．为此利用 $\{\boldsymbol{q}_0,\ \boldsymbol{q}_1,\ \cdots,\ \boldsymbol{q}_{i-1},\ A\boldsymbol{q}_{i-1}\}=\{\boldsymbol{r}_0,\ A\boldsymbol{r}_0,\ \cdots,\ A^i\boldsymbol{r}_0\}$，于是可以将 $\boldsymbol{q}_i$ 表示成

$$\boldsymbol{q}_i=A\boldsymbol{q}_{i-1}+\lambda_0^{(i-1)}\boldsymbol{q}_0+\lambda_1^{(i-1)}\boldsymbol{q}_1+\cdots+\lambda_{i-1}^{(i-1)}\boldsymbol{q}_{i-1},$$

利用 $[\boldsymbol{q}_i,\ \boldsymbol{q}_j]=0$，$j=0$、1、2、…、$i-1$，知道

$$\begin{aligned}\lambda_j^{(i-1)}&=-[A\boldsymbol{q}_{i-1},\ \boldsymbol{q}_j]/[\boldsymbol{q}_j,\ \boldsymbol{q}_j]\\&=-[\boldsymbol{q}_{i-1},\ A\boldsymbol{q}_j]/[\boldsymbol{q}_j,\ \boldsymbol{q}_j],\end{aligned}$$

但 $A\boldsymbol{q}_j=\boldsymbol{q}_{j+1}-\sum\limits_{l=0}^{j}\lambda_l^{(j)}\boldsymbol{q}_l$，

$$\lambda_j^{(i-1)}=-\left[\boldsymbol{q}_{i-1},\ \boldsymbol{q}_{j+1}-\sum_{=0}^{j}\lambda_l^{(j)}\boldsymbol{q}_l\right]\Big/[\boldsymbol{q}_j,\ \boldsymbol{q}_j],$$

可知当 $j+1<i-1$ 或 $j<i-2$ 时

$$\lambda_j^{(i-1)}=0,$$

由此
$$\boldsymbol{q}_i=A\boldsymbol{q}_{i-1}+\lambda_{i-1}^{(i-1)}\boldsymbol{q}_{i-1}+\lambda_{i-2}^{(i-1)}\boldsymbol{q}_{i-2},$$

并且
$$\begin{aligned}\lambda_{i-1}^{(i-1)}&=-[A\boldsymbol{q}_{i-1},\ \boldsymbol{q}_{i-1}]/[\boldsymbol{q}_{i-1},\ \boldsymbol{q}_{i-1}],\\\lambda_{i-2}^{(i-1)}&=-[\boldsymbol{q}_{i-1},\ \boldsymbol{q}_{i-1}]/[\boldsymbol{q}_{i-2},\ \boldsymbol{q}_{i-2}],\end{aligned}$$

这样每次计算 $\boldsymbol{q}_i$ 只要用到 $\boldsymbol{q}_{i-1}$、$\boldsymbol{q}_{i-2}$ 两个向量，节约了存贮量和计算量．

新的计算程式如下：

取一个初始向量 $\boldsymbol{x}_0$，计算 $\boldsymbol{r}_0=A\boldsymbol{x}_0-\boldsymbol{b}$，取 $\boldsymbol{q}_0=\boldsymbol{r}_0$，

$$\left\{\begin{array}{l}\alpha_k=-(\boldsymbol{r}_k,\ \boldsymbol{q}_k)/(A\boldsymbol{q}_k,\ \boldsymbol{q}_k),\\\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\alpha_k\boldsymbol{q}_k,\ \boldsymbol{r}_{k+1}=\boldsymbol{r}_k+\alpha_kA\boldsymbol{q}_k,\\\boldsymbol{q}_{k+1}=A\boldsymbol{q}_k+\lambda_k^{(k)}\boldsymbol{q}_k+\lambda_{k-1}^{(k)}\boldsymbol{q}_{k-1},\\\lambda_k^{(k)}=-(A\boldsymbol{q}_k,\ A\boldsymbol{q}_k)/(A\boldsymbol{q}_k,\ \boldsymbol{q}_k),\\\lambda_{k-1}^{(k)}=-(A\boldsymbol{q}_k,\ \boldsymbol{q}_k)/(A\boldsymbol{q}_{k-1},\ \boldsymbol{q}_{k-1}),\ \lambda_{-1}^{(0)}=0,\\k=0、1、2、\cdots.\end{array}\right.\tag{18}$$

共轭斜量法产生的序列 $\{\boldsymbol{x}_k\}$，有一个重要的关系式

$$(\boldsymbol{r}_k,\ \boldsymbol{r}_j)=0,\quad j=0、1、\cdots、k-1, \tag{19}$$

实际上

$$(\boldsymbol{r}_1,\ \boldsymbol{r}_0)=(\boldsymbol{r}_0+\alpha_0 A\boldsymbol{q}_0,\ \boldsymbol{r}_0)=(\boldsymbol{r}_0,\ \boldsymbol{r}_0)+\alpha_0(A\boldsymbol{q}_0,\ \boldsymbol{r}_0)$$
$$=(\boldsymbol{r}_0,\ \boldsymbol{r}_0)-(\boldsymbol{r}_0,\ \boldsymbol{r}_0)=0.$$

由 $\boldsymbol{x}_k$ 的定义知

$$[\boldsymbol{x}_k-\tilde{\boldsymbol{x}},\ \boldsymbol{q}_j]=0,\quad j=0,\ 1,\ \cdots,\ k-1,$$

即 $$(\boldsymbol{r}_k,\ \boldsymbol{q}_j)=0,\quad j=0,\ 1,\ \cdots,\ k-1.$$

只要证明 $\boldsymbol{r}_j\in\{\boldsymbol{q}_0,\ \boldsymbol{q}_1,\ \cdots,\ \boldsymbol{q}_j\}$，(19) 式就得到证明了．实际上

$$\boldsymbol{r}_0=\boldsymbol{q}_0,$$

假设 $\boldsymbol{r}_l\in\{\boldsymbol{q}_0,\ \boldsymbol{q}_1,\ \cdots,\ \boldsymbol{q}_l\}$ 对于 $l=0,\ 1,\cdots,j-1$ 都成立，来证明 $\boldsymbol{r}_j\in\{\boldsymbol{q}_0,\ \boldsymbol{q}_1,\ \cdots,\ \boldsymbol{q}_j\}$．因为

$$\boldsymbol{r}_j=\boldsymbol{r}_{j-1}+\alpha_{j-1}A\boldsymbol{q}_{j-1}$$
$$=\boldsymbol{r}_{j-1}+\alpha_{j-1}(\boldsymbol{q}_j-\lambda_{j-1}^{(j-1)}\boldsymbol{q}_{j-1}-\lambda_{j-2}^{(j-1)}\boldsymbol{q}_{j-2}),$$

因此 $\boldsymbol{r}_j\in\{\boldsymbol{q}_0,\ \boldsymbol{q}_1,\ \cdots,\ \boldsymbol{q}_j\}$，从而证明了(19)．

如果 $\boldsymbol{r}_j=0$，那么 $\boldsymbol{x}_j=\tilde{\boldsymbol{x}}$；因为 n 维空间最多有 n 个非零彼此正交向量，因此总有一个 $j\leqslant n$ 使 $\boldsymbol{r}_j=0$．这也就是说共轭斜量法，在理论上是有限步就可以结束的．

如果 $\boldsymbol{r}_0、\boldsymbol{r}_1、\cdots、\boldsymbol{r}_j$ 全不为零，则可以证明

$$\{\boldsymbol{r}_0,\ \boldsymbol{r}_1,\ \cdots,\ \boldsymbol{r}_j\}=\{\boldsymbol{q}_0,\ \boldsymbol{q}_1,\ \cdots,\ \boldsymbol{q}_j\}$$
$$=\{\boldsymbol{r}_0,\ A\boldsymbol{r}_0,\ \cdots,\ A^j\boldsymbol{r}_0\},$$

从而 $\boldsymbol{r}_0、\boldsymbol{r}_1、\cdots、\boldsymbol{r}_j$ 是 $\{\boldsymbol{r}_0,\ A\boldsymbol{r}_0,\ \cdots,\ A^j\boldsymbol{r}_0\}$ 的一组正交基．前面已经证明了 $\boldsymbol{r}_i\in\{\boldsymbol{q}_0,\ \boldsymbol{q}_1,\ \cdots,\ \boldsymbol{q}_j\}$，$i=0,\ 1,\ 2,\ \cdots,\ j$，现在只要证明 $\boldsymbol{q}_i\in\{\boldsymbol{r}_0,\ \boldsymbol{r}_1,\ \cdots,\ \boldsymbol{r}_j\}$，$i=0,\ 1,\ 2,\ \cdots,\ j$，那么就有

$$\{\boldsymbol{r}_0,\ \boldsymbol{r}_1,\ \cdots,\ \boldsymbol{r}_j\}=\{\boldsymbol{q}_0,\ \boldsymbol{q}_1,\ \cdots,\ \boldsymbol{q}_j\}.$$

实际上从

$$\boldsymbol{r}_i=\boldsymbol{r}_{i-1}+\alpha_{i-1}(\boldsymbol{q}_i-\lambda_{i-1}^{(i-1)}\boldsymbol{q}_{i-1}-\lambda_{i-2}^{(i-1)}\boldsymbol{q}_{i-2}),\quad i=1,\ \cdots,\ j.$$

知道 $\alpha_{i-1}\neq0$, $i=1,\ \cdots,\ j$, 否则

$$\boldsymbol{r}_i=\boldsymbol{r}_{i-1}.$$

由 $(\boldsymbol{r}_i,\ \boldsymbol{r}_{i-1})=0$ 得 $(\boldsymbol{r}_i,\ \boldsymbol{r}_i)=0$, 与 r_0、r、$\cdots$、r_j 不全为零矛盾. 由 $\alpha_{i-1}\neq0$, 从而可知 $\boldsymbol{q}_i$ 可由 $\boldsymbol{r}_i$、$\boldsymbol{r}_{i-1}$、$\cdots$、$\boldsymbol{r}_0$ 线性组合来表示, 即 $\boldsymbol{q}_i\in\{\boldsymbol{r}_0,\ \boldsymbol{r}_1,\ \cdots,\ \boldsymbol{r}_j\}$, $i=0,\ 1,\ 2,\ \cdots,\ j$.

现在我们可以利用 $\boldsymbol{r}_i$ 的正交性, 来进一步减少 A-正交化过程的计算量和存贮量. 在(18)的计算过程中, 在计算得到 $\boldsymbol{r}_{k+1}$ 后, 若 $\boldsymbol{r}_{k+1}\neq0$ 利用下式计算 $\boldsymbol{q}_{k+1}$,

$$\boldsymbol{q}_{k+1}=\boldsymbol{r}_{k+1}+\omega_0^{(k)}\boldsymbol{q}_0+\omega_1^{(k)}\boldsymbol{q}_1+\cdots+\omega_k^{(k)}\boldsymbol{q}_k.$$

由 $[\boldsymbol{q}_{k+1},\ \boldsymbol{q}_j]=0$, $j=0,\ 1,\ \cdots,\ k$, 知

$$\omega_j^{(k)}=-[\boldsymbol{r}_{k+1},\ \boldsymbol{q}_j]/[\boldsymbol{q}_j,\ \boldsymbol{q}_j],$$

$$[\boldsymbol{r}_{k+1},\ \boldsymbol{q}_j]=(\boldsymbol{r}_{k+1},\ A\boldsymbol{q}_j)=\left(\boldsymbol{r}_{k+1},\ \frac{\boldsymbol{r}_{j+1}-\boldsymbol{r}_j}{\alpha_j}\right),$$

因此当 $j=0,\ 1,\cdots,\ k-1$ 时, $\omega_j^{(k)}=0$, 这样

$$\boldsymbol{q}_{k+1}=\boldsymbol{r}_{k+1}+\omega_k^{(k)}\boldsymbol{q}_k,$$

$\boldsymbol{q}_0,\ \boldsymbol{q}_1,\cdots,\ \boldsymbol{q}_k$, $k+1$ 个向量中, 只有 $\boldsymbol{q}_k$ 在计算 $\boldsymbol{q}_{k+1}$ 时要用到, 又减少了一个向量. 此时

$$\begin{aligned}\omega_k^{(k)}&=-[\boldsymbol{r}_{k+1},\ \boldsymbol{q}_k]/[\boldsymbol{q}_k,\ \boldsymbol{q}_k]\\&=-(\boldsymbol{r}_{k+1},\ \boldsymbol{r}_{k+1})/\alpha_k[\boldsymbol{q}_k,\ \boldsymbol{q}_k],\end{aligned}$$

$$\begin{aligned}\alpha_k&=-(\boldsymbol{r}_k,\ \boldsymbol{q}_k)/(A\boldsymbol{q}_k,\ \boldsymbol{q}_k)\\&=-(\boldsymbol{r}_k,\ \boldsymbol{r}_k+\omega_{k-1}^{(k-1)}\boldsymbol{q}_{k-1})/(A\boldsymbol{q}_k,\ \boldsymbol{q}_k)\\&=-(\boldsymbol{r}_k,\ \boldsymbol{r}_k)/(A\boldsymbol{q}_k,\ \boldsymbol{q}_k),\end{aligned}$$

于是

$$\omega_k^{(k)}=(\boldsymbol{r}_{k+1},\ \boldsymbol{r}_{k+1})/(\boldsymbol{r}_k,\ \boldsymbol{r}_k).$$

将 $\omega_k^{(k)}$ 记为 λ_k, 我们得到另外一个计算程式:

取一个初始向量 x_0, 计算 $\boldsymbol{r}_0=A\boldsymbol{x}_0-\boldsymbol{b}$, 取 $\boldsymbol{q}_0=\boldsymbol{r}_0$,

$$
\begin{cases}
\alpha_k = -(\boldsymbol{r}_k, \boldsymbol{r}_k)/(A\boldsymbol{q}_k, \boldsymbol{q}_k), \\
\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{q}_k, \ \boldsymbol{r}_{k+1} = \boldsymbol{r}_k + \alpha_k A\boldsymbol{q}_k, \\
\boldsymbol{q}_{k+1} = \boldsymbol{r}_{k+1} + \lambda_k \boldsymbol{q}_k, \\
\lambda_k = (\boldsymbol{r}_{k+1}, \boldsymbol{r}_{k+1})/(\boldsymbol{r}_k, \boldsymbol{r}_k), \\
k = 0, 1, 2, \cdots.
\end{cases} \tag{20}
$$

(20)式也是流行的共轭斜量法的计算程式.

我们介绍了三种求 A-正交化向量组的方法即(16),(18)和(20),各种方法所求的 $\boldsymbol{q}_k$,长度 $\|\boldsymbol{q}_k\|$ 可能会有差别,但它们之间最多差一个常数因子. 从计算量和存贮量节省来说(20)是最好了. 使用(20)式计算 $\boldsymbol{x}_{k+1}$,需要用到 $\boldsymbol{x}_k$、$\boldsymbol{r}_k$、$\boldsymbol{q}_k$、$A\boldsymbol{q}_k$,一般它们都要存放在快速存贮器中,这样要占用 $4n$ 个存贮单元,这一点比起其他求方程组的解的迭代法来说是比较费的. 如果在每一步迭代中,多计算一次 $A\boldsymbol{q}_k$,就可以不存放 $A\boldsymbol{q}_k$,即在计算

$$\alpha_k = -(\boldsymbol{r}_k, \boldsymbol{r}_k)/(A\boldsymbol{q}_k, \boldsymbol{q}_k)$$

中,计算一次 $A\boldsymbol{q}_k$,作内积 $(A\boldsymbol{q}_k, \boldsymbol{q}_k)$ 后,不保留 $A\boldsymbol{q}_k$,在计算

$$\boldsymbol{r}_{k+1} = \boldsymbol{r}_k + \alpha_k A\boldsymbol{q}_k$$

时,再计算一次 $A\boldsymbol{q}_k$,计算完 $\boldsymbol{r}_k$ 后,也不保留 $A\boldsymbol{q}_k$,这是一种增加计算量换得减少存贮量的措施.

共轭斜量法不但可用来求解系数矩阵为对称正定时的方程组,而且还可以用来求解系数矩阵为对称非负定时的方程组,也即系数矩阵 A 除了有正的特征值外,还有零特征值时,即行列式 $\det(A)=0$ 时,也可使用共轭斜量法. 考虑方程组

$$A\boldsymbol{x} = \boldsymbol{b} \tag{1}$$

的系数矩阵 A 的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 中有 $\lambda_1, \lambda_2, \cdots, \lambda_l$ 是零而 $\lambda_{l+1}>0$,λ_i 对应的特征向量为 y_i. 于是(1)要有解的充要条件是

$$(\boldsymbol{b}, \boldsymbol{y}_i)=0, \quad i=1, 2, \cdots, l.$$

记集合

$$H_0=\{\boldsymbol{x} \mid (\boldsymbol{x}, \boldsymbol{y}_i)=0, i=1, 2, \cdots, l\},$$

显然 H_0 是一个线性子空间，并且 $\boldsymbol{b}\in H_0$. 在 H_0 上引进函数 $(A\boldsymbol{x}, \boldsymbol{y})$, $\boldsymbol{x}, \boldsymbol{y}\in H_0$, 容易验证 $(A\boldsymbol{x}, \boldsymbol{y})$ 在 H_0 上满足内积的所有条件，因此可以看作空间 H_0 上的内积.

因为(1)的解非唯一，可以差一个零特征向量的线性组合 $\sum_{i=1}^{l}\beta_i\boldsymbol{y}_i$. 因此对于任意取定的初始向量 $\boldsymbol{x}_0$, 有一个解 $\tilde{\boldsymbol{x}}$, 使得 $\boldsymbol{x}_0-\tilde{\boldsymbol{x}}\in H_0$, 自然有 $\boldsymbol{r}_0=A\boldsymbol{x}_0-\boldsymbol{b}\in H_0$, $A\boldsymbol{r}_0\in H_0$, $\cdots$, $A^k\boldsymbol{r}_0\in H_0$, 将在 H_0 上的内积 $(A\boldsymbol{x}, \boldsymbol{y})$ 也记为 $[\boldsymbol{x}, \boldsymbol{y}]$, 同样可以考虑多步斜量法，使

$$\left[\boldsymbol{x}_0+\sum_{j=0}^{k-1}\beta_jA^j\boldsymbol{r}_0-\tilde{\boldsymbol{x}}, \boldsymbol{x}_0+\sum_{j=0}^{k-1}\beta_jA^j\boldsymbol{r}_0-\tilde{\boldsymbol{x}}\right]$$

达到极小，同样可以将 $\boldsymbol{r}_0$, $A\boldsymbol{r}_0$, $\cdots$, $A^k\boldsymbol{r}_0$ A-正交化，就得到共轭斜量法，计算程式与(20)一样，并且有收敛速度的估计

$$\|\tilde{\boldsymbol{x}}-\boldsymbol{x}_k\|\leqslant\sqrt{\frac{\lambda_n}{\lambda_{l+1}}}\widetilde{E}_k\|\tilde{\boldsymbol{x}}-\boldsymbol{x}_0\|,$$

其中

$$\widetilde{E}_k=\left[T_k\left(\frac{\lambda_n+\lambda_{l+1}}{\lambda_n-\lambda_{l+1}}\right)\right]^{-1}.$$

在[5]中介绍了在实际问题中使用共轭斜量法解系数阵非负定时的方程组的一个例子.

当系数矩阵对称不定的情况，即 A 有正的特征值，也有负的特征值时，一般不能再用(20)求解了. 对于这样的方程组，用什么办法求解，也是目前大家较关心的问题，在第2章§5将介绍一种 Lanczos 算法. 当然在那种情况的各种算法的理论，也没有象对称正定时共轭斜量法那样完全，因此也是进一步需要研究的问题.

§4 不完全分解、预处理共轭斜量法

上一节介绍的共轭斜量法，与其它求解方程组的方法比较起来，有下列几个很重要的优点:

1．每步迭代需要计算的量，用到系数矩阵 A 的．只有 A 与向量的乘法 $A\boldsymbol{q}$，因此可以充分利用 A 的稀疏性．

2．不要预先估计别的参数就可以计算，这一点不象契贝谢夫半迭代法、超松弛法等方法．

3．每次迭代所需要计算，都是向量之间的运算，可充分为第四代计算机(向量运算计算机)所利用．

但是实际计算中，使用共轭斜量法，也会碰到问题，主要是收敛慢的问题，一方面是舍入误差的影响，更主要的是系数矩阵 A 的条件数太大．在 §1 中我们已经讲过条件数 $p=\frac{\lambda_n}{\lambda_1}$，

$\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}=\frac{p+1}{p-1}=\frac{1+\frac{1}{p}}{1-\frac{1}{p}}$，而共轭斜量法的收敛速度快慢依赖于

$$\left[T_k\left(\frac{\lambda_n+\lambda_1}{\lambda_n-\lambda_1}\right)\right]^{-1}=\left[T_k\left(\frac{1+\frac{1}{p}}{1-\frac{1}{p}}\right)\right]^{-1}$$

趋于 0 的速度．根据契贝谢夫多项式的定义，可知

$$\begin{aligned}T_k\left(\frac{1+\varepsilon}{1-\varepsilon}\right)&=\frac{1}{2}\left\{\left(\frac{1+\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}\right)^k+\left(\frac{1+\sqrt{\varepsilon}}{1-\sqrt{\varepsilon}}\right)^{-k}\right\}\\&\approx\frac{1}{2}\exp(2k\sqrt{\varepsilon}),\quad \text{当 } k\sqrt{\varepsilon}>1,\ \varepsilon\to 0.\end{aligned}$$

当 p 很大时，$1/p$、$\sqrt{1/p}$ 都很小，因此 $\exp(-2k\sqrt{1/p})$ 收敛于 0 很慢.

很早以前就有人提出，对于方程组

$$A\boldsymbol{x}=\boldsymbol{b} \tag{1}$$

作等价变换，变成

$$B\boldsymbol{x}=\boldsymbol{f}, \tag{21}$$

或者变成

$$B\boldsymbol{y}=\boldsymbol{g}, \tag{22}$$

只要 $\boldsymbol{x}$ 能够容易从 $\boldsymbol{y}$ 获得，而 B 仍然保持对称正定，并且 B 的条件数 $p(B)$ 要比 A 的条件数 $p(A)$ 小，这样解方程(21)或(22)要比解方程(1)有利. 当然使用共轭斜量法于(21)、(22)收敛速度也会比对于(1)的要快. 这就是预处理(Precondition)的思想. 但是很多预处理，尽管条件数改善了，但乃不象 A 那样具有稀疏性，并且增加很多预处理的计算量，这可参见[6].

1977 年 J. A. Meijerink 和 A. Van der Vorst [7] 提出一种称为不完全分解(Incomplete decomposition)的办法，将 A 分解成

$$A=LL^T+R,$$

这里 L 是下三角阵；一方面使 LL^T 尽可能接近 A，另一方面使 L 保持跟 A 一样的稀疏性，或者具有其他形状的稀疏性.

完全分解是对矩阵 A 进行三角分解 $A=LL^T$，不完全分解是对矩阵 $A-R$ 进行三角分解 LL^T. 对于不完全分解，因为有矩阵 R，可以变化，因此 L 中那些元素为 0，可以预先规定，不过这种规定也不能完全任意，还须考虑到 LL^T 接近 A，但是这是不易检验的，因此实际计算时，就考虑使 R 有较多零元素.

为了明白具体做法，举 4×4 对称正定矩阵 A 的不完全分解为例：$A=(a_{ij})$，$R=(r_{ij})$，要求分解成 LL^T，其中 L 的形状如下：

$$L=\begin{pmatrix} l_{11} & & & 0 \\ l_{21} & l_{22} & & \\ 0 & l_{32} & l_{33} & \\ 0 & 0 & l_{43} & l_{44} \end{pmatrix},$$

即 $l_{31}=l_{41}=l_{42}=0$，比较

$$LL^T=(A-R),$$

有 $l_{11}^2=a_{11}-r_{11}$，因为 A 正定，故 $a_{11}>0$，因此可取 $r_{11}=0$，于是 $l_{11}=\sqrt{a_{11}}$；再由 $l_{11}l_{21}=(a_{12}-r_{12})$，可以取 $r_{12}=0$，$l_{21}=a_{12}/l_{11}$。由 $l_{11}l_{31}=(a_{31}-r_{31})$，因为 $l_{31}=0$，因此必须取 $r_{31}=a_{31}$；同样由 $l_{41}=0$，从 $l_{11}l_{41}=(a_{41}-r_{41})$ 确定 $r_{41}=a_{41}$。然后由

$$l_{21}^2+l_{22}^2=a_{22}-r_{22},$$

$$l_{22}^2=a_{22}-l_{21}^2-r_{22},$$

如果 $a_{22}-l_{21}^2>0$，那么可取 $r_{22}=0$，在完全分解时，当 A 对称正定必有 $a_{22}-l_{21}^2>0$，对于不完全分解，就不一定能保证 $a_{22}-l_{21}^2>0$，如果 $a_{22}-l_{21}^2\leqslant0$，取一个绝对值小的负数 r_{22}，使 $a_{22}-l_{21}^2-r_{22}>0$，这样

$$l_{22}=\sqrt{a_{22}-l_{21}^2-r_{22}},$$

由 $l_{21}l_{31}+l_{22}l_{32}=(a_{32}-r_{32})$，可取 $r_{32}=0$，$l_{32}=a_{32}/l_{22}$，由 $l_{31}l_{41}+l_{22}l_{42}=a_{42}-r_{42}$，要 $l_{42}=0$，必须 $r_{42}=a_{42}$，由 $l_{31}^2+l_{32}^2+l_{33}^2=a_{33}-r_{33}$，$l_{33}^2=a_{33}-l_{32}^2-r_{33}$，同样如果 $a_{33}-l_{32}^2>0$，那么 $r_{33}=0$，否则取 r_{33} 为一个使 $a_{33}-l_{32}^2-r_{33}>0$ 的负数，于是 $l_{33}=\sqrt{a_{33}-l_{32}^2-r_{33}}$，由 $l_{31}l_{41}+l_{32}l_{42}+l_{33}l_{43}=a_{33}-r_{43}$，取 $r_{43}=0$，$l_{43}=a_{43}/l_{33}$；最后由

$$l_{44}^2=a_{44}-l_{43}^2-r_{44}$$

决定 r_{44} 和 $l_{44}=\sqrt{a_{44}-l_{43}^2-r_{44}}$. 按此得到的

$$R=\begin{pmatrix} 0 & 0 & a_{31} & a_{41} \\ 0 & r_{22} & 0 & a_{42} \\ a_{31} & 0 & r_{33} & 0 \\ a_{41} & a_{42} & 0 & r_{44} \end{pmatrix},$$

如果 A 中的元素 a_{31}, a_{41}, a_{42} 有的是 0, 那么 R 中就会有更多的零. 由此可以设想, A 稀疏时, R 会有很多零, 对于对角元素 r_{22}, r_{33}, r_{44}, 一般不能保证为 0, 但是文献[7]中证明, 当 A 是对称正定的 M-阵时, 即

$a_{ij}\leqslant 0\,(i\neq j)$, $a_{ii}>0$, A^{-1} 元素全为正时, 不管 L 的形状如何, 所有 $r_{ii}=0$.

对于 A 不是 M-阵时, 不能保证所有 $r_{ii}=0$, 例如

$$A=\begin{pmatrix} 3 & -2 & 0 & 2 \\ -2 & 3 & -2 & 0 \\ 0 & -2 & 3 & -2 \\ 2 & 0 & -2 & 3 \end{pmatrix},$$

要求 L 与 A 有相同的稀疏性, 即 $l_{31}=l_{42}=0$, 此时

$$L=\begin{pmatrix} \sqrt{3} & & & \text{\Large 0} \\ -2/\sqrt{3} & \sqrt{5}/3 & & \\ 0 & -2\sqrt{3}/\sqrt{5} & \sqrt{3}/\sqrt{5} & \\ 2/\sqrt{3} & 0 & -2\sqrt{5}/\sqrt{3} & a \end{pmatrix},$$

$$R=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4/3 \\ 0 & 0 & 0 & 0 \\ 0 & 4/3 & 0 & -(5+a^2) \end{pmatrix}.$$

所谓不完全分解预处理共轭斜量法，即是先作 A 的某种不完全分解

$$A=LL^T+R,$$

再将方程组 $$A\boldsymbol{x}=\boldsymbol{b}$$

化成 $$L^{-1}AL^{-T}L^T\boldsymbol{x}=L^{-1}\boldsymbol{b},$$

然后对方程组

$$L^{-1}AL^{-T}\boldsymbol{y}=L^{-1}\boldsymbol{b} \tag{23}$$

使用共轭斜量法，求得 $\boldsymbol{y}_k$，再由 $L^Tx=\boldsymbol{y}_k$ 求得 $\boldsymbol{x}_k$. 因为 A 对称正定，易知 $L^{-1}AL^{-T}$ 也是对称正定，因此对(23)使用共轭斜量法，是可行的.

具体计算的程式如下：

先对方程组(1)的 A 进行不完全分解，得 L.

取 $\boldsymbol{x}_0$，算 $\boldsymbol{r}_0=A\boldsymbol{x}_0-\boldsymbol{b}$，求 $\boldsymbol{p}_0=(LL^T)^{-1}\boldsymbol{r}_0$,

$$\alpha_k=-((LL^T)^{-1}\boldsymbol{r}_k,\ \boldsymbol{r}_k)/(A\boldsymbol{p}_k,\ \boldsymbol{p}_k),$$

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\alpha_k\boldsymbol{p}_k,$$

$$\boldsymbol{r}_{k+1}=\boldsymbol{r}_k+\alpha_kA\boldsymbol{p}_k,$$

$$\lambda_k=((LL^T)^{-1}\boldsymbol{r}_{k+1},\ \boldsymbol{r}_{k+1})/((LL^T)^{-1}\boldsymbol{r}_k,\ \boldsymbol{r}_k),$$

$$\boldsymbol{p}_{k+1}=(LL^T)^{-1}\boldsymbol{r}_{k+1}+\lambda_k\boldsymbol{p}_k,$$

$$k=0,\ 1,\ 2,\ \cdots,$$

以上就是不完全分解、预处理共轭斜量法的计算程式，这个方法简记为 ICCG 法.

程式中出现 $(LL^T)^{-1}r_k$，表示要解一个方程组

$$(LL^T)y=r_k,$$

但因为 L 是下三角阵，又是稀疏的，因此计算量和存贮量都可以不大.

有很多实际计算的报告，指出这一方法非常成功，可参见[7]、[8]，例如[8]中指出对于一个激光聚变(Laser fusion)方

程

$$\frac{\partial f}{\partial t}=\nabla D\nabla f+r(f^e-f),$$

这里 D、r、f^e 是给定的非负函数. 边界条件是

$$\boldsymbol{n}\cdot\nabla f=0 \quad 和 \quad f=0$$

两部分的混合.

用差分方法解此方程,用隐式 5 点差分格式,共取 555 个节点,化成一个线性方程组

$$A\boldsymbol{x}=\boldsymbol{b},$$

系数矩阵 A 的条件数 $p=2\times10^{10}$.

如果记 $\tilde{x}$ 为正确解,要求计算到

$$\varepsilon=\|\tilde{\boldsymbol{x}}-\boldsymbol{x}\|/\|\tilde{\boldsymbol{x}}\|=10^{-6},$$

用下列五种方法进行计算,所要求的迭代次数比较如下:

1. ICCG 方法

$$A=LL^T+R,$$

其中 L 的稀疏性跟 A 同,记为 ICCG(0),需 25 次迭代.

2. 取最佳松弛因子的块超松弛法,需 765 次迭代.

3. 隐式交替方向法 1,需 10200 次迭代.

4. 隐式交替方向法 2,需 4750 次迭代.

5. Gauss-Seidel 法,需 208000 次迭代.

对于这个问题,ICCG(0) 比 Gauss-Seidet 法要快了 8000 倍.

这个例子说明 ICCG 法在实用中有很惊人的成绩. 但是迄今对这一方法的理论分析,还没有重要结果. 究竟 $L^{-1}AL^{-T}$ 的条件数比 A 的条件数,一般情况下小多少? 当 L 取怎样的稀疏性时较好? 当 A 不是 M-阵时, l_{ii} 取什么数较好? 这些问题都没有解决. 目前有很多学者,正在研究 ICCG 方法.

第 2 章

对称三对角矩阵

对称三对角矩阵，是一种运算，存放都比较简单的矩阵．一个对称矩阵可以通过有限步计算，正交相似变换成一个对称三对角矩阵，因此很多求对称矩阵特征值的方法，第一步先把矩阵化成对称三对角阵，然后再来求三对角阵的特征值．有些求解系数矩阵对称的线性代数方程组的方法，也是先把系数矩阵，归化成对称三对角阵，然后再来解系数矩阵为对称三对角阵的方程组．因此在矩阵计算中，对称三对角矩阵，已成为一种有用的工具．在这一章介绍对称三对角矩阵的性质，一方面为了后面两章的应用，另外方面也有它本身的独立意义．另外还介绍了解对称线性方程组的Lanczos算法．

§1 Jacobi 矩阵

设

$$A=\begin{pmatrix} b_1 & c_1 & & & \\ a_1 & b_2 & c_2 & & \\ & a_2 & b_3 & c_3 & \\ & & \ddots & \ddots & \ddots \\ & & & & c_{n-1} \\ & & & a_{n-1} & b_n \end{pmatrix}$$

称为三对角矩阵，如果 a_i、b_i、c_i 都是实数，且 $a_ic_i>0$ 则称 A 为 Jacobi 矩阵．

从它的形状可知，它的 k 阶顺序主子阵也是三对角阵，记为 A_k，显然 $A_1=b_1$, $A_n=A$.

设 A_k 的特征多项式为 $\varphi_k(\lambda)$，则

$$\varphi_k(\lambda)=\det(\lambda I-A_k)$$

$$=\begin{vmatrix} \lambda-b_1 & -c_1 & & & 0 \\ -a_1 & \lambda-b_2 & -c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & & & -c_{k-1} \\ 0 & & & -a_{k-1} & \lambda-b_k \end{vmatrix}$$

$$=(\lambda-b_k)\varphi_{k-1}(\lambda)$$

$$+c_{k-1}\begin{vmatrix} \lambda-b_1 & -c_1 & & 0 \\ -a_1 & \lambda-b_2 & -c_2 & \\ & \ddots & \ddots & \ddots \\ & -a_{k-3} & \lambda-b_{k-2} & -c_{k-2} \\ 0 & & 0 & -a_{k-1} \end{vmatrix}$$

$$=(\lambda-b_k)\varphi_{k-1}(\lambda)-a_{k-1}c_{k-1}\varphi_{k-2}(\lambda).$$

令 $\varphi_0(\lambda)=1$，所以对 $k=2, 3, \cdots, n$ 成立关系式

$$\varphi_k(\lambda)=(\lambda-b_k)\varphi_{k-1}(\lambda)-a_{k-1}c_{k-1}\varphi_{k-2}(\lambda). \tag{1}$$

定理2.1 Jacobi 阵的特征多项式序列

$$\varphi_n(\lambda)、\varphi_{n-1}(\lambda)、\cdots、\varphi_1(\lambda)、\varphi_0(\lambda) \tag{2}$$

是任何区间$[a, b]$内的 Sturm 序列(见[9]).

证明

1. 序列最后一个多项式为 $\varphi_0(\lambda)=1$，没有根.

2. 序列中相邻两个多项式 $\varphi_k(\lambda)$, $\varphi_{k-1}(\lambda)$ 没有公根. 否则由递推公式(1)知，这个根也是 $\varphi_{k-2}(\lambda)$ 的根，从而可以推

知它也是 $\varphi_0(\lambda)$ 的根，此与 $\varphi_0(\lambda)$ 无根矛盾.

3. 若 λ_0 是 $\varphi_{k-1}(\lambda)$ 的根，则 $\varphi_k(\lambda_0)$ 与 $\varphi_{k-2}(\lambda_0)$ 反号. 这是因为 $\varphi_k(\lambda_0)=-c_{k-1}a_{k-1}\varphi_{k-2}(\lambda_0)$，而 $c_{k-1}a_{k-1}>0$ 故 $\varphi_k(\lambda_0)$ 与 $\varphi_{k-2}(\lambda_0)$ 反号.

序列(2)具有上述三个性质，因此为任何区间$[a, b]$内的一个 Sturm 序列. 证毕.

现在来分析序列(2)的零点分布.

1. $\varphi_0(\lambda)=1$，说明 $\varphi_0(\lambda)$ 在整个实数轴$(-\infty, \infty)$上都保持正号.

2. $\varphi_1(\lambda)=\lambda-b_1$，有一个零点，即 $x_1^{(1)}=b_1$，$\varphi_1(\lambda)$ 在 $\lambda<b_1$ 时为负，在 $\lambda>b_1$ 时为正.

3. $\varphi_2(\lambda)=(\lambda-b_2)\varphi_1(\lambda)-a_1c_1\varphi_0(\lambda)$ 是一个二次多项式，首项系数为正，因此在 $\lambda\to\pm\infty$ 时，$\varphi_2(\lambda)$ 为正，但在 $\lambda=x_1^{(1)}$ 处 $\varphi_2(x_1^{(1)})=-a_1c_1$ 为负，这说明 $\varphi_2(\lambda)$ 有二个零点 $x_1^{(2)}$、$x_2^{(2)}$，并且

$$x_1^{(2)}<x_1^{(1)}<x_2^{(2)}.$$

4. 一般地说，可以用数学归纳法证明如下性质：$\varphi_{k-1}(\lambda)$ 有 $k-1$ 个实单根 $x_1^{(k-1)}$、$x_2^{(k-1)}$、…、$x_{k-1}^{(k-1)}$，$\varphi_k(\lambda)$ 有 k 个实单根 $x_1^{(k)}$、$x_2^{(k)}$、…、$x_k^{(k)}$，并成立下列关系

$$x_1^{(k)}<x_1^{(k-1)}<x_2^{(k)}<x_2^{(k-1)}<\cdots<x_{k-1}^{(k)}<x_{k-1}^{(k-1)}<x_k^{(k)},$$

实际上，假如这一性质对 k 成立，我们证明这一性质对 $k+1$ 也成立.

考察 $\varphi_{k-1}(\lambda)$在$(x_i^{(k)}-\varepsilon, x_i^{(k)}+\varepsilon)$中的符号，这里 ε 是充分小的正数.

当 k 是奇数时，$k-1$ 是偶数，$\varphi_{k-1}(\lambda)$在$(-\infty, x_1^{(k-1)})$中是正的，$\varphi_{k-1}(\lambda)$在$(x_1^{(k-1)}, x_2^{(k-1)})$中是负的，一般可知 $\varphi_{k-1}(\lambda)$

在$(x_{i-1}^{(k-1)}, x_i^{(k-1)})$中的符号同$(-1)^{i-1}$. 因为$(x_i^{(k)}-\varepsilon, x_i^{(k)}+\varepsilon)$ $\subset(x_{i-1}^{(k-1)}, x_i^{(k-1)})$，因此当$k$是奇数时$\varphi_{k-1}(\lambda)$在$(x_i^{(k)}-\varepsilon, x_i^{(k)}+\varepsilon)$中的符号同$(-1)^{i-1}$.

当k是偶数时，$\varphi_{k-1}(\lambda)$在$(-\infty, x_1^{(k-1)})$中的符号为负，在$(x_1^{(k-1)}, x_2^{(k-1)})$中符号为正，一般，在$(x_{i-1}^{(k-1)}, x_i^{(k-1)})$中的符号同$(-1)^i$，因此$\varphi_{k-1}(\lambda)$在$(x_i^{(k)}-\varepsilon, x_i^{(k)}+\varepsilon)$中的符号同$(-1)^i$.

因为$\varphi_{k+1}(x_i^{(k)})$的符号与$\varphi_{k-1}(x_i^{(k)})$的相反，因此当k是奇数时，$\varphi_{k+1}(x_i^{(k)})$的符号同$(-1)^i$，这样$\varphi_{k+1}(\lambda)$在$(x_i^{(k)}, x_{i+1}^{(k)})$中至少有一个根，于是在$(x_1^{(k)}, x_k^{(k)})$中至少有$k-1$个根，因为$k$是奇数，$\varphi_{k+1}(x_1^{(k)})$是负的，$\varphi_{k+1}(x_k^{(k)})$也是负的，而$\varphi_{k+1}(\lambda)$是偶数次多项式，在$\lambda\to\pm\infty$处都是正的，因此在$(-\infty, x_1^{(k)})$中至少有一个根，在$(x_k^{(k)}, \infty)$中也至少有一个根，$\varphi_{k+1}(\lambda)$至多有$k+1$个根，于是知道在$(x_i^{(k)}, x_{i+1}^{(k)})$中有且只有一个根. 记为$x_{i+1}^{(k+1)}$，同样在$(-\infty, x_1^{(k)})$中有且只有一个根记为$x_1^{(k+1)}$，在$(x_k^{(k)}, \infty)$中有且只有一个根记为$x_{k+1}^{(k+1)}$. 这就证明了当$k$是奇数时所述性质对$k+1$也成立.

对于k是偶数的情况类似的可以证明. 我们有

定理 2.2 Jacobi 矩阵的特征多项式序列$\varphi_0(\lambda)$、$\varphi_1(\lambda)$、…、$\varphi_n(\lambda)$中任意一个多项式$\varphi_k(\lambda)$，有k个实单根$x_1^{(k)}$、$x_2^{(k)}$、…、$x_k^{(k)}$，并且它们与$\varphi_{k-1}(\lambda)$的$k-1$个实单根$x_1^{(k-1)}$、$x_2^{(k-1)}$、…、$x_{k-1}^{(k-1)}$，有如下隔离性质

$$x_1^{(k)}<x_1^{(k-1)}<x_2^{(k)}<x_2^{(k-1)}<\cdots<x_{k-1}^{(k-1)}<x_k^{(k)}.$$

§2 对称三对角矩阵的唯一归化定理

形如

$$
T=\begin{pmatrix}
\alpha_1 & \beta_1 & & & 0 \\
\beta_1 & \alpha_2 & \beta_2 & & \\
 & \ddots & \ddots & \ddots & \\
 & & \ddots & \ddots & \beta_{n-1} \\
0 & & & \beta_{n-1} & \alpha_n
\end{pmatrix}
$$

的方阵称为对称三对角矩阵，当 $\beta_i \neq 0\ (i=1,\ 2,\ \cdots,\ n-1)$ 时，T 称为不可约。显然它是特殊的 Jacobi 阵，因此关于 Jacobi 阵的结论，它都具有。如果有 $\beta_i=0$，那么矩阵是可约的 [9, p. 230]．可约对称三对角阵的方程组求解问题和特征值问题，都可化成几个低阶的不可约矩阵的问题来讨论．本章以后几节将进一步介绍对称三对角矩阵的重要性质．

首先指出，任意一个实对称矩阵，都可以通过正交相似变换，变成对称三对角矩阵．

取 $A=(a_{ij})$ 是任意一个实对称矩阵，它可以通过 Householder 镜像变换 Q_1 使

$$
Q_1AQ_1^T=\begin{pmatrix}
a_{11} & \tau & 0 & \cdots & 0 \\
\tau & & & & \\
0 & & & & \\
\vdots & & A_2 & & \\
0 & & & &
\end{pmatrix},
$$

其中 $Q_1=I-2\boldsymbol{w}_1\boldsymbol{w}_1^T=Q_1^T$,

$$
\boldsymbol{w}_1^T=(0,\ a_{21}-\tau,\ a_{31},\ \cdots,\ a_{n1})/h,
$$

$$
\tau=-(\operatorname{sign} a_{21})\sqrt{\sum_{j=2}^{n} a_{j1}^2}\ ,\quad h=\sqrt{2\tau(\tau-a_{21})}\ ,
$$

熟知 Q_1 是正交变换．τ 之所以要取成与 $-a_{21}$ 同号，目的是使 h 尽可能大一些，有利于计算的稳定．如果记

$$A_2=\begin{pmatrix} a'_{22} & a'_{23} & \cdots & a'_{2n} \\ a'_{32} & a'_{33} & \cdots & a'_{3n} \\ \cdots & \cdots & \cdots & \cdots \\ a'_{n2} & a'_{n3} & \cdots & a'_{nn} \end{pmatrix},$$

则再通过镜像变换

$$Q_2=I-2\boldsymbol{w}_2\boldsymbol{w}_2^T,$$

$\boldsymbol{w}_2$ 是由 A_2 的第一列构成

$$\boldsymbol{w}_2^T=(0,\ 0,\ a'_{32}-\tau',\ a'_{42},\ \cdots,\ a'_{n2})/h',$$

其中 $\tau'=-(\operatorname{sign} a'_{32})\sqrt{\sum_{j=2}^{n}(a'_{j2})^2},\quad h'=\sqrt{2\tau'(\tau'-a'_{32})}.$

$$Q_2Q_1AQ_1^TQ_2^T=\begin{pmatrix} a_{11} & \tau & 0 & 0 & \cdots & 0 \\ \tau & a'_{22} & \tau' & 0 & \cdots & 0 \\ 0 & \tau' & & & & \\ 0 & 0 & & & & \\ \vdots & \vdots & & A_3 & & \\ 0 & 0 & & & & \end{pmatrix},$$

Q_2Q_1 仍是正交变换. 依此类推, 经过 Q_1、Q_2、$\cdots$、Q_{n-2} 镜像变换, 可以将 A 相似变换成一个对称三对角矩阵.

也可以通过平面旋转变换, 将对称矩阵 A, 正交相似变换成对称三对角矩阵. 记

$$T_{l,k}=\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & c & & & s & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & -s & & & c & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{pmatrix}\begin{matrix} \\ \\ \\ l\text{行} \\ \\ \\ \\ k\text{行} \\ \\ \\ \\ \end{matrix}$$

$$\qquad\qquad l\text{列}\qquad\qquad k\text{列}$$

$s^2+c^2=1$,称为(l, k)平面上的旋转矩阵,它也是一个正交阵.

$$B=(b_{ij})=T_{lk}AT_{lk}^T,$$

称为将 A 通过(l, k)平面旋转相似变换成 B, 此时 B 的元素与 A 的元素之间的关系是

$$\begin{aligned}
&b_{li}=b_{il}=a_{il}c+a_{ik}s,\\
&b_{ki}=b_{ik}=-a_{il}s+a_{ik}c, \quad i\neq l,\ k;\\
&b_{ll}=a_{ll}c^2+2a_{lk}sc+a_{kk}s^2,\\
&b_{kk}=a_{ll}s^2-2a_{lk}sc+a_{kk}c^2,\\
&b_{lk}=b_{kl}=(a_{kk}-a_{ll})sc+a_{lk}(c^2-s^2),
\end{aligned}$$

其余的 $b_{ij}=a_{ij}$(当 $i\neq l$、k, 且 $j\neq l$、k).

于是可知 B 仍是对称, 跟 A 相比只改变二行二列元素. 如果 $i\neq l$、k, $a_{il}=a_{ik}=0$, 那么通过(l, k)平面的旋转后,仍有 $b_{il}=b_{ik}=0$; 如果 $a_{il}^2+a_{ik}^2\neq 0$, 则取 $s=a_{ik}/\sqrt{a_{il}^2+a_{ik}^2}$, $c=a_{il}/\sqrt{a_{il}^2+a_{ik}^2}$, 可使 $b_{ik}=0$.

利用上述性质,提出两种方案将一个实对称矩阵,正交相似变换成一个对称三对角矩阵.

1. 当矩阵是满秩矩阵时. 将下列足标的元素,依次逐个化成零,

$$\begin{aligned}
&(1,\ n),\ (1,\ n-1),\ \cdots,\ (1,\ 4),\ (1,\ 3),\\
&(2,\ n),\ (2,\ n-1),\ \cdots,\ (2,\ 4),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&(n-2,\ n),
\end{aligned}\tag{3}$$

而在化足标如(i, j)的元素为 0 时,采用的旋转矩阵为$T_{j-1,j}$, 即在$(j-1, j)$平面上作旋转, 如果当时的矩阵为 $\widetilde{A}=(\tilde{a}_{pq})$, 则取

$$s=\tilde{a}_{ij}/\sqrt{\tilde{a}_{ij-1}^2+\tilde{a}_{ij}^2}, \quad c=\tilde{a}_{ij-1}/\sqrt{\tilde{a}_{ij-1}^2+\tilde{a}_{ij}^2},$$

如果 $\tilde{a}_{ij}=0$, 那么 $T_{j-1,j}=I$, 此时这个旋转可以不做. 这样至

多进行了 $\frac{(n-1)(n-2)}{2}$ 次旋转相似变换后，就将 A 化成一个对称三对角矩阵．即

$$T_{n-1,n}\cdots T_{n-2,n-1}T_{n-1,n}AT_{n-1,n}^{T}T_{n-2,n-1}^{T}\cdots T_{n-1,n}^{T}=H$$

是一个对称三对角矩阵，上述表示式中，同样记号 $T_{j-1,j}$ 会出现不止一次，例如 $T_{n-1,n}$ 出现 $n-2$ 次，但每次出现的 $T_{n-1,n}$，它的 s、c 可能是不同的，其它的 T_{ij} 也是这样．

这个过程比起镜像变换的办法来，计算量可能要大一倍．因此对于满秩矩阵的情况，一般使用镜像变换的办法．

2．当矩阵是带形时．矩阵元素 a_{ij}，当 $j>i$ 时称它在第 $j-i$ 条超对角线(Superdiagonal)上，a_{1n} 在第 $n-1$ 条超对角线上，a_{12} 在第1条超对角线上．所谓带形对称矩阵就是存在一个自然数 $m<n$，对任何 $k>m$，第 k 条超对角线上的元素全为0．当然，有兴趣的是 $m\ll n$ 的情况．对于这样的 m，称矩阵为带宽是 $2m+1$ 的带形矩阵．

对于这种带形矩阵，通过平面旋转，逐个把所有三对角线外的元素化成零．化的次序是先对第 m 条超对角线上的所有元素，将它们全部化成零后，再对第 $m-1$ 条超对角线上的所有元素，最后对第2条超对角线上的元素．在化每条超对角线上元素为零时，所取元素的次序，是按照行的次序；例如在化第 k 条次对角线上元素时，所取元素的足标次序为 $(1,\ k+1)$，$(2,\ k+2)$，$\cdots$，$(n-k,\ n)$．

在消去足标为 (i,j) 的元素时，使用的平面旋转为 $T_{j-1,j}$，设此时的矩阵为 $\tilde{A}=(\tilde{a}_{pq})$，则取

$$s=\tilde{a}_{ij}/\sqrt{\tilde{a}_{ij-1}^{2}+\tilde{a}_{ij}^{2}},$$

$$c=\tilde{a}_{ij-1}/\sqrt{\tilde{a}_{ij-1}^{2}+\tilde{a}_{ij}^{2}},$$

如果 $\tilde{a}_{ij}=0$，此时 $T_{j-1,j}=I$，不必进行这一次旋转。相似变换

后的矩阵

$$B=T_{j-1,j}\tilde{A}T_{j-1,j}^{T}=(b_{pq}),$$

就有 $b_{ij}=0$.

不过必须特别指出，在消去元素 $\tilde{a}_{ij}$ 时，次序大于 $j-i$ 的超对角线上元素，已经化成零了．即 $\tilde{A}$ 中的元素 $\tilde{a}_{l,k}$ 当 $k-l>j-i$ 时，$\tilde{a}_{l,k}=0$，但是对于元素 $\tilde{a}_{j,2j-i}$，如果 $2j-i\leqslant n$，矩阵中是有这个元素的，它是在第 $j-i$ 条超对角线上的，因为它的行足标 $j>i$，因此它可能不为 0．这样 B 的元素

$$b_{j-1,2j-i}=\tilde{a}_{j-1,2j-i}c+\tilde{a}_{j,2j-i}s=\tilde{a}_{j,2j-i}s\neq 0.$$

$\tilde{a}_{j-1,2j-i}$ 是在第 $j-i+1$ 条超对角线上，所以为零．于是在第 $j-i+1$ 条超对角线上产生了新的非零元 $b_{j-1,2j-i}$．尽管产生了新的非零元，但是要注意，这个新的非零元的列足标比被消的对象 $\tilde{a}_{ij}$ 的列足标要大 $j-i$．这一点使人们想到再用 $T_{2j-i-1,2j-i}$ 将 $b_{j-1,2j-1}$ 化成零．即得到

$$A=T_{2j-i-1,2j-i}BT_{2j-i-1,2j-i}^{T}=(\tilde{\tilde{a}}_{pq}),$$

此时取

$$s=b_{j-1,2j-i}/\sqrt{b_{j-1,2j-i}^2+b_{j-1,2j-i-1}^2},$$

$$c=b_{j-1,2j-i-1}/\sqrt{b_{j-1,2j-i}^2+b_{j-1,2j-i-1}^2},$$

就有

$$\tilde{\tilde{a}}_{j-1,2j-i}=0.$$

同样如果 $2j-i+(j-i)=3j-2i\leqslant n$，而 $b_{2j-i,3j-2i}$ 又不为零，则又产生一个新的非零元 $\tilde{\tilde{a}}_{2j-i-1,\ 3j-2i}$，不过列的足标又增加 $j-i$；这样产生一个新的非零元，用平面旋转将它化成零，新的非零元又产生，不过列足标增大 $j-i$，因为列足标最大是 n．因此当 $\left[\frac{n-j}{j-i}\right]$ 步后，就不会再有新的非零元产生．这样的办法，消去一个 $\tilde{a}_{ij}$，总共要化 $\left[\frac{n-j}{j-i}\right]+1$ 次平面旋转变换．但这样的过程，最大的优点是保持带宽，矩阵的带宽以外部分不必占用存贮单元．而镜象变换，一般就不能保持带宽，在这

一点上比镜像变换优越. 很多实际问题产生的矩阵，都是稀疏带形矩阵，因此这个办法是很有实用价值的. 这个方法最早由 H. R. Schwarz 给出，见[24]. 图 2 给出这个过程的示意.

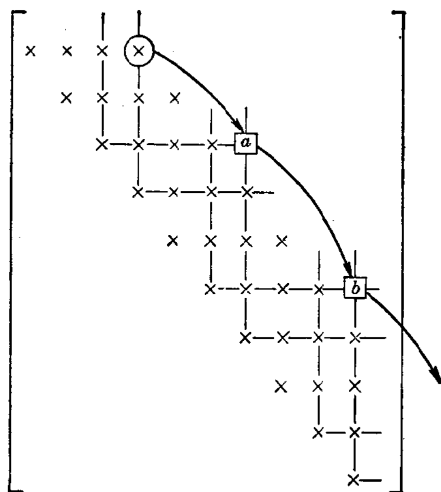

图 2

从上面介绍知道，一个对称矩阵 A，通过正交变换相似于一个对称三对角矩阵，可以有各种途径，最后得到的相似的对称三对角阵，也可能不相同. 但是有下面定理.

定理 2.3 对于 $n\times n$ 实对称矩阵 A，如果有正交矩阵 Q，使得 $Q^TAQ=T$ 是一个对称三对角矩阵，并且 T 的第一条超对角线上的元素全为正的，那么 Q 和 T 完全由 Q 的第一列 $\boldsymbol{q}_1$ 和 A 所决定，或者完全由 Q 的第 n 列 $\boldsymbol{q}_n$ 和 A 所决定.

证明 记

$$T=\begin{pmatrix}\alpha_1 & \beta_1 & & & 0\\ \beta_1 & \alpha_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1}\\ 0 & & & \beta_{n-1} & \alpha_n\end{pmatrix},$$

$$Q=(\boldsymbol{q}_1,\ \boldsymbol{q}_2,\ \cdots,\ \boldsymbol{q}_n),$$

本定理要证明 α_i, β_i, $\boldsymbol{q}_i$, 由 $\boldsymbol{q}_1$ 和 A 唯一确定.

由 $Q^TAQ=T$ 得 $AQ=QT$ 比较此式两边的各列,有

$$A\boldsymbol{q}_1=\alpha\cdot\boldsymbol{q}_1+\beta\cdot\boldsymbol{q}_2, \tag{4}$$

$$A\boldsymbol{q}_i=\alpha_i\boldsymbol{q}_i+\beta_i\boldsymbol{q}_{i+1}+\beta_{i-1}\boldsymbol{q}_{i-1},\quad i=2,\ 3,\ \cdots,\ n, \tag{5}$$

由 Q 是正交阵,它的各列彼此正交,因此从(4)知

$$\alpha_1=\boldsymbol{q}_1^TA\boldsymbol{q}_1,$$

$$\beta_1\boldsymbol{q}_2=A\boldsymbol{q}_1-\alpha_1\boldsymbol{q}_1,$$

利用 $\beta_1>0$ 和 $\|\boldsymbol{q}_2\|=1$, 故

$$\beta_1=\|A\boldsymbol{q}_1-\alpha_1\boldsymbol{q}_1\|,$$

从而可知,α_1、β_1 完全由 A 和 $\boldsymbol{q}_1$ 确定. 因为 $\beta_1>0$, 故

$$\boldsymbol{q}_2=(A\boldsymbol{q}_1-\alpha_1\boldsymbol{q}_1)/\beta_1,$$

也由 A 和 $\boldsymbol{q}_1$ 完全确定. 一般地如果: $\boldsymbol{q}_2$, $\boldsymbol{q}_3$, $\cdots$, $\boldsymbol{q}_{i-1}$, $\boldsymbol{q}_i$; α_1, α_2, $\cdots$, α_{i-1}; β_1, β_2, $\cdots$, β_{i-1} 已经由 A、$\boldsymbol{q}_1$ 完全确定, 那么利用(5),知 $\alpha_i=\boldsymbol{q}_i^TA\boldsymbol{q}_i$, 若 $i\neq n$ 有

$$\beta_i\boldsymbol{q}_{i+1}=A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1},$$

$$\beta_i=\|A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1}\|,$$

$$\boldsymbol{q}_{i+1}=(A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1})/\beta_i,$$

完全被 $\boldsymbol{q}_i$、$\boldsymbol{q}_{i-1}$ 和 A 所确定, 从而被 $\boldsymbol{q}_1$ 和 A 所确定. 同理可知所有 $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\cdots$, $\boldsymbol{q}_{n-1}$; α_1, $\cdots$, α_n; β_1, $\cdots$, β_{n-1} 完全可由 A 和 $\boldsymbol{q}_n$ 所确定. 证毕.

定理 2.3 中 T 的要求是: 第 1 条超对角线上元素为正. 如果 T 的第 1 条超对角线上元素有正有负, 但不为零, 是否

也有类似的结论．为了回答这个问题先看

引理 2.1 若

$$T=\begin{pmatrix} \alpha_1 & \beta_1 & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & & & \beta_{n-1} \\ 0 & & & \beta_{n-1} & \alpha_n \end{pmatrix},\quad \beta_i\neq 0,\ i=1,\ 2,\ \cdots,\ n-1,$$

则存在非奇异对角阵 $\Delta=\mathrm{diag}(\delta_1,\ \delta_2,\cdots,\ \delta_n)$，其中 $\delta_1=1$, $\delta_{i+1}=\mathrm{sign}(\beta_i\delta_i)$, $i=1$、2、…、$n-1$, 使得

$$\Delta T\Delta^{-1}=\begin{pmatrix} \alpha_1 & |\beta_1| & & & 0 \\ |\beta_1| & \alpha_2 & |\beta_2| & & \\ & |\beta_2| & \ddots & \ddots & \\ & & \ddots & \ddots & |\beta_{n-1}| \\ 0 & & & |\beta_{n-1}| & \alpha_n \end{pmatrix}. \tag{6}$$

证明

$\Delta T\Delta^{-1}$

$$=\begin{pmatrix} \alpha_1 & (\beta_1\delta_1/\delta_2) & & & 0 \\ (\beta_1\delta_2/\delta_1) & \alpha_2 & (\beta_2\delta_2/\delta_3) & & \\ & (\beta_2\delta_3/\delta_2) & \ddots & \ddots & \\ & & & \ddots & (\beta_{n-1}\delta_{n-1}/\delta_n) \\ 0 & & & (\beta_{n-1}\delta_n/\delta_{n-1}) & \alpha_n \end{pmatrix},$$

因为 $\qquad \beta_k\delta_{k+1}/\delta_k=\beta_k\delta_k/\delta_{k+1}=|\beta_k|,$

故(6)成立．证毕．

定理 2.4 对于 $n\times n$ 实对称阵 A，如果有正交阵 $P=[\boldsymbol{p}_1,\ \boldsymbol{p}_2,\ \cdots,\ \boldsymbol{p}_n]$，使得

$$P^TAP=T=\begin{pmatrix}\alpha_1 & \beta_1 & & 0\\ \beta_1 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_{n-1}\\ 0 & & \beta_{n-1} & \alpha_n\end{pmatrix}$$

是一个对称三对角矩阵，并且 T 的第 1 条超对角线元素不为 0，则 T 和 $\boldsymbol{p}_2, \boldsymbol{p}_3, \cdots, \boldsymbol{p}_n$ 完全由 A 和 $\boldsymbol{p}_1$，以及 T 的第 1 条超对角线上元素的符号所完全确定.

证明 如果所有 $\beta_i>0$，则由定理 2.3 知结果成立. 如果 β_i 中有正有负，则取引理 2.1 中对角阵 Δ，使

$$\Delta T\Delta=T_+=\begin{pmatrix}\alpha_1 & |\beta_1| & & & 0\\ |\beta_1| & \alpha_2 & |\beta_2| & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & |\beta_{n-1}|\\ 0 & & & |\beta_{n-1}| & \alpha_n\end{pmatrix},$$

于是 $$\Delta P^TAP\Delta^{-1}=\Delta T\Delta^{-1}=T_+,$$

记 $Q=P\Delta^{-1}=(\boldsymbol{q}_1\boldsymbol{q}_2\cdots\boldsymbol{q}_n)=\left(\boldsymbol{p}_1, \frac{1}{\delta_2}\boldsymbol{p}_2, \frac{1}{\delta_3}\boldsymbol{p}_3, \cdots, \frac{1}{\delta_n}\boldsymbol{p}_n\right)$,

$Q^TQ=\Delta^{-1}P^TP\Delta^{-1}=\Delta^{-1}\Delta^{-1}=I$，故 Q 是一个正交阵. 由定理 2.3 知 Q 的各列和 T_+ 完全被 A 和 $\boldsymbol{q}_1$ 所确定，而 $\boldsymbol{q}_1=\boldsymbol{p}_1$，因此 Q 的各列和 T_+ 完全被 A 和 $\boldsymbol{p}_1$ 所确定. 知道了 β_i 的符号，就可以从 $|\beta_i|$ 得到 β_i，从 $\delta_i\boldsymbol{q}_i$ 得到 $\boldsymbol{p}_i$；故 A 和 $\boldsymbol{p}_1$ 完全确定 P 和 T. 证毕.

本节所介绍的通过镜像相似变换，将 A 化成对称三对角矩阵，是将 A 左乘和右乘一系列初等阵 $Q_i=I-2\boldsymbol{w}_i\boldsymbol{w}_i^T$ 后，相似成一个对称三对角阵 T_1，即

$$T_1=Q_sQ_{s-1}\cdots Q_2Q_1AQ_1Q_2\cdots Q_s,$$

$Q_1Q_2\cdots Q_s$ 是一个正交阵. 注意到 $\boldsymbol{w}_i$ 的第 1 个分量总是零, 因此 Q_i 的第 1 列为 $\boldsymbol{e}_1=(1, 0, \cdots, 0)^T$, 第 1 行为 $\boldsymbol{e}_1^T=(1, 0, 0, \cdots, 0)$, 即

$$Q_i=\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & & & & \\ \vdots & & H_i & & \\ 0 & & & & \end{pmatrix},$$

因此 $Q=Q_1Q_2\cdots Q_s$ 的第 1 列为 $\boldsymbol{e}_1$; 如果镜像变换结果的第一条超对角线上元素 τ 都不为 0, 那么 T_1 和 Q 完全由 A 和 $\boldsymbol{e}_1$ 以及那些 τ 的符号所决定. 因此这个过程与定理 2.3 中的下述过程产生的结果 T 和 Q 是一样的:

取
$$\boldsymbol{q}_1=\boldsymbol{e}_1, \quad \beta_0=0,$$
$$\alpha_i=(A\boldsymbol{q}_i, \boldsymbol{q}_i),$$
$$\beta_i\boldsymbol{q}_{i+1}=A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1},$$
$$\beta_i=\pm\|A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1}\|$$
(符号由对应 τ 的符号确定),
$$\boldsymbol{q}_{i+1}=(A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1})/\beta_i,$$
$$i=1, 2, \cdots, n-1,$$
$$\alpha_n=(A\boldsymbol{q}_n, \boldsymbol{q}_n).$$

由此我们也可以知道, 镜像变换过程中, 出现 $\tau=0$ 的充分必要条件是 n 个向量

$$\boldsymbol{e}_1, A\boldsymbol{e}_1, \cdots, A^{n-1}\boldsymbol{e}_1$$

线性相关. 这是因为 $\beta_i=0$, 充要条件是 $A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1}=0$, 而后者意味着

$$\boldsymbol{e}_1, A\boldsymbol{e}_1, \cdots, A^{i-1}\boldsymbol{e}_1, A^i\boldsymbol{e}_1$$

线性相关.

再来考察平面旋转变换将 A 化成对称三对角矩阵的过

程．记第 k 次旋转时，使用的 s、c 为 s_k、c_k，$t=\frac{(n-1)(n-2)}{2}$．设 A 是满秩矩阵，按足标次序(3)，通过平面旋转，将 A 相似于对称三对角阵 T_2，于是

$$T_2=T_{n-1,n}(s_t,\ c_t)\cdots T_{n-2,n-1}(s_2,\ c_2)T_{n-1,n}(s_1,\ c_1)A$$
$$T^T_{n-1,n}(s_1,\ c_1)T^T_{n-2,n-1}(s_2,\ c_2)\cdots T^T_{n-1,n}(s_t,\ c_t),$$

记 $P=T^T_{n-1,n}(s_1,\ c_1)T^T_{n-2,n-1}(s_2,\ c_2)\cdots T^T_{n-1,n}(s_t,\ c_t)$，在这些旋转矩阵 $T^T_{j-1,j}$ 中，都有 $j\geqslant 3$，因此 $T^T_{j-1,j}(s_k,\ c_k)$ 的第 1 行第 1 列都与单位阵 I 的第 1 行第 1 列相同．这样也可知道 P 的第 1 列为 $\boldsymbol{e}_1$，于是可知若 T_1 和 T_2 的第 1 条超对角线上的元素全为正，则 $T_2=T_1$，$P=Q$．如果 T_2 的第 1 条超对角线上的元素有正有负，但不为零，则 $T_1=\Delta T_2\Delta^{-1}$，$Q=P\Delta^{-1}$，这里 $\Delta=\mathrm{diag}(1,\ \delta_2,\ \cdots,\ \delta_n)$ 是象引理 2.1 中所给的 Δ 那样的矩阵．

上面指出镜像变换和旋转变换，都是 $\boldsymbol{q}_1=\boldsymbol{e}_1$ 的正交相似变换

$$Q^TAQ=T,$$

将 A 变成对称三对角阵 T．当然也可以取任意单位向量 $\boldsymbol{a}$ 为 $\boldsymbol{q}_1$，此时进行的过程为

$$\begin{cases}\boldsymbol{q}_1=\boldsymbol{a},\\ \alpha_i=(A\boldsymbol{q}_i,\ \boldsymbol{q}_i),\quad \beta_i\boldsymbol{q}_{i+1}=A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1},\\ \beta_i=\|A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1}\|,\\ \boldsymbol{q}_{i+1}=(A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1})/\beta_i,\\ i=1,\ 2,\ \cdots,\ n-1,\end{cases}\tag{7}$$

(7)称为 Lanczos 过程．

如果在过程中某个 $\beta_i=0$，则可取任何与 $\boldsymbol{q}_1,\ \boldsymbol{q}_2,\ \cdots,\ \boldsymbol{q}_i$ 都正交的单位向量作为 $\boldsymbol{q}_{i+1}$，过程可继续进行．

§3 对称三对角矩阵的极值性质*

记对称三对角矩阵

$$T=\begin{pmatrix} \alpha_1 & \beta_1 & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & & & \beta_{n-1} \\ 0 & & & \beta_{n-1} & \alpha_n \end{pmatrix}$$

的截段

$$\begin{pmatrix} \alpha_l & \beta_l & & & 0 \\ \beta_l & \alpha_{l+1} & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & & \beta_{m-1} \\ 0 & & & \beta_{m-1} & \alpha_m \end{pmatrix}$$

为 $T_{l,m}$, $1\leqslant l<m\leqslant n$; 它的特征多项式为 $\chi_{l,m}(\lambda)$, 因此 $T_{1,n}=T$. 为了方便起见记 $T_j=T_{1,j}$, 特征多项式为 $\chi_j(\lambda)$.

引理 2.2 若 $l<j$ 则 $T_j^l\boldsymbol{e}_1$ 的第 $l+1$ 个分量为 $\beta_1\beta_2\cdots\beta_l$, 而第 $l+2$ 个分量至第 j 个分量都为 0, 这里 $\boldsymbol{e}_1$ 是 j 维单位向量 $(1,\ 0,\ \cdots,\ 0)^T$.

证明

$$T_j e_1=\begin{pmatrix} \alpha_1 \\ \beta_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

* 本节引理 2.2, 定理 2.5 的结果取自[10].

第2个分量为β_1，第3个分量至第j个分量都为0，因此命题对$l=1$成立．今假设命题对$l=k<j-1$成立．

$T_j^k e_1$的第$k+1$个分量为$\beta_1\beta_2\cdots\beta_k$，而第$k+2$个分量至第$j$个分量都为0．考察$T_j T_j^k e_1$的第$k+2$个元素．因为$T_j$的第$k+2$行为三个元素$\beta_{k+1}$，$\alpha_{k+2}$，$\beta_{k+2}$分别在第$k+1$列，第$k+2$列，第$k+3$列，因此$T_j T_j^k e_1$的第$k+2$个元素为$\beta_1\beta_2\cdots\cdot\beta_k\beta_{k+1}$．而$T_j$的第$k+3$行的三个元素$\beta_{k+2}$，$\alpha_{k+3}$，$\beta_{k+3}$是在第$k+2$列，$k+3$列，$k+4$列，因此$T_j T_j^k e_1$的第$k+3$个元素为0，同样可以证明次序数大于$k+3$的元素也为0．证毕．

定理2.5 设T的第1条超对角线上的元素全为正，则对$j=1, 2, \cdots, n-1$成立

a．$\beta_1\beta_2\cdots\beta_j=\|\chi_j(T)e_1\|=\min\limits_{\psi\in mP_j}\|\psi(T)e_1\|$，

这里mP_j表示全体首项系数为1的j次多项式的集合．

b．$e_{j+1}=\chi_j(T)e_1/\beta_1\beta_2\cdots\beta_j$．

证明 由引理2.2知对任意$\psi(\lambda)\in mP_j$，$\psi(T)$的第$j+1$个分量为$\beta_1\beta_2\cdots\beta_j$，因此

$$\beta_1\beta_2\cdots\beta_j\leqslant\|\psi(T)e_1\|,$$

特别取mP_j中的$\chi_j(\lambda)$，证明

$$\chi_j(T)e_1=\beta_1\beta_2\cdots\beta_j e_{j+1}.$$

记$E_j=(e_1, e_2, \cdots, e_j)$是$n\times j$矩阵．首先证明下列命题

$$E_j^* T^l e_1=T_j^l e_1$$

对$l\leqslant j$成立．实际上此命题意思即为，$T^l e_1$的为首j个分量与$T_j^l e_1$的分量完全相同．对于$l=1$，$j=1$这命题显然成立，对于$l=1$，$j\geqslant 2$

$$Te_1=\begin{pmatrix}\alpha_1\\ \beta_1\\ 0\\ \vdots\\ 0\end{pmatrix}_{1\times n}, \quad E_j^*Te_1=\begin{pmatrix}\alpha_1\\ \beta_1\\ 0\\ \vdots\\ 0\end{pmatrix}_{1\times j}=T_je_1,$$

因此命题对 $l=1$ 成立. 假定它对 $k\leqslant j-1$ 也成立. 考虑 $T^{k+1}e_1=T(T^ke_1)$, 因为 T^ke_1 的第 $k+2$ 个分量开始后面的分量全为 0, 因此第 $j+1$ 个分量以及后面的分量都为 0, 这样 $T(T^ke_1)$ 的为首 j 个分量与 $T_j(T_j^ke_1)$ 的相同, 命题得证.

现在来讨论 $\chi_j(T)e_1$, 由引理 2.2 知道它的第 $j+1$ 个分量为 $\beta_1\beta_2\cdots\beta_j$, 另外 $\chi_j(T)e_1$ 的为首 j 个分量与 $\chi_j(T_j)e_1$ 的 j 个分量相同, 而后者根据 Cayley-Hamilton 定理(见[9]中 p. 363)知 $\chi_j(T_j)=0$, 即 $\chi_j(T_j)e_1=0$, 因此 $\chi_j(T)e_1$ 除了第 $j+1$ 个分量外全为 0, 即

$$\chi_j(T)e_1=\beta_1\beta_2\cdots\beta_je_{j+1}. \text{证毕.}$$

推论 设 $Q=[q_1q_2\cdots q_n]$ 是一个正交矩阵, 又设 A 是实对称矩阵. 若 $Q^*AQ=T$ 是一个对称三对角矩阵, $\beta_i>0$, $i=1, 2, \cdots, n-1$, 则

$$\beta_1\beta_2\cdots\beta_j=\|\chi_j(A)q_1\|=\min_{\psi\in mP_j}\|\psi(A)q_1\| \tag{8}$$

且
$$q_{j+1}=\chi_j(A)q_1/\beta_1\beta_2\cdots\beta_j. \tag{9}$$

证明 由 $Q^*AQ=T$, 因此 $Q^*\chi_j(A)Q=\chi_j(T)$, 故

$$\chi_j(T)e_1=Q^*\chi_j(A)Qe_1=Q^*\chi_j(A)q_1, \tag{10}$$

由定理 2.5

$$\begin{aligned}\beta_1\beta_2\cdots\beta_j&=\|\chi_j(T)e_1\|=\|\chi_j(A)q_1\|\\ &\leqslant\min_{\psi\in mP_j}\|\psi(T)e_1\|=\min_{\psi\in mP_j}\|\psi(A)q_1\|,\end{aligned}$$

(8)式得到证明, 另外从(10)式知

$$\beta_1\beta_2\cdots\beta_j \boldsymbol{e}_{j+1} = Q^* \chi_j(A) \boldsymbol{q}_1,$$

两边左乘 Q, 即得(9)式, 从而本推论证毕.

在对称方程组求解问题, 和对称矩阵特征值问题中, 我们常碰到如下的线性方程组

$$T_j \boldsymbol{f}_j = \alpha \boldsymbol{e}_1,$$

这里 α 是常数, $f_j = (\delta_1^{(j)}, \delta^{(j)}, \cdots, \delta_j^{(j)})^T$ 是未知向量, 要求将 $\delta_i^{(j)}$ 用 T_j 的元素和 α 表示出来. 有如下定理(参见[31])

定理 2.6 若 $\det(T_j) = d_j \neq 0$, 则

$$\delta_k^{(j)} = (-1)^{k-1} \alpha \frac{\beta_1\beta_2\cdots\beta_{k-1} \det(T_{k+1,j})}{\det(T_j)}, \tag{11}$$

$$k = 1, 2, \cdots, j, \quad 这里 \quad \det(T_{j+1,j}) = 1.$$

证明 利用 Cramer 法则

$$\delta_k^{(j)} = s_k / d_j, \tag{12}$$

这里

第 k 列

$$s_k = \det \begin{pmatrix} \alpha_1 & \beta_1 & & & \alpha & & & & \\ \beta_1 & \alpha_2 & \ddots & & \vdots & & & \mathbf{0} & \\ & \ddots & \ddots & & 0 & & & & \\ & & \ddots & \beta_{k-2} & 0 & & & & \\ & & \beta_{k-2} & \alpha_{k-1} & 0 & & & & \\ & & & \beta_{k-1} & 0 & \beta_k & & & \\ & & & & 0 & \alpha_{k+1} & \beta_{k+2} & & \\ & \mathbf{0} & & & \vdots & \beta_{k+1} & \alpha_{k+2} & \ddots & \\ & & & & \vdots & & \ddots & \ddots & \beta_{j-1} \\ & & & & 0 & & & \beta_{j-1} & \alpha_j \end{pmatrix},$$

将 s_k 按第 k 列展开,

$$s_k=(-1)^{k-1}\alpha$$

$$\times\det\begin{pmatrix}
\beta_1 & \alpha_2 & \beta_2 & & & & & & \\
 & \beta_2 & \alpha_3 & \ddots & & & & 0 & \\
 & & \ddots & \ddots & \beta_{k-2} & & & & \\
 & & & \ddots & \alpha_{k-1} & & & & \\
 & & & & \beta_{k-1} & \beta_k & & & \\
 & & & & & \alpha_{k+1} & \beta_{k+1} & & \\
 & & & & & \beta_{k+1} & \alpha_{k+2} & \ddots & \\
 & 0 & & & & & \ddots & \ddots & \beta_{j-1} \\
 & & & & & & & \beta_{j-1} & \alpha_j
\end{pmatrix},$$

上述表示式中右端的行列式可以看作是一个块上三角矩阵的行列式. 利用块上三角矩阵的行列式等于对角块行列式之积的性质(见[9]中p.144), 就得

$$s_k=(-1)^{k-1}\alpha$$

$$\times\det\begin{pmatrix}
\beta_1 & \alpha_2 & & \\
 & \beta_2 & \ddots & \\
 & & \ddots & \alpha_{k-1} \\
0 & & & \beta_{k-1}
\end{pmatrix}\det\begin{pmatrix}
\alpha_{k+1} & \beta_{k+1} & & 0 \\
\beta_{k+1} & \alpha_{k+2} & \ddots & \\
 & \ddots & \ddots & \beta_{j-1} \\
0 & & \beta_{j-1} & \alpha_j
\end{pmatrix}$$

$$=(-1)^{k-1}\alpha\beta_1\beta_2\cdots\beta_{k-1}\det(T_{k+1,j}),$$

将此结果代入(12)即得(11). 证毕.

§4 Thompson–McEnteggert–Paige 公式和特征值反问题

记 adj(B) 为矩阵 B 的伴随阵, 即

$$B\,\mathrm{adj}(B)=\det(B)I,$$

1968 年 Thompson 和 McEnteggert[12] 给出如下的特征向量与伴随阵之间的关系：

定理 2.7 设 $n\times n$ 实对称矩阵 A 的 n 个标准正交特征向量为 $z_1, z_2, \cdots, z_n$，它们对应的特征值依次为 λ_1、λ_2、$\cdots$、λ_n，$\chi(\lambda)$ 为 A 的特征多项式 $\det(\lambda I-A)$，则

$$\operatorname{adj}(\lambda_i I-A)=\chi'(\lambda_i)z_i z_i^*.$$

证明 取常数 $\mu\neq\lambda_j$，$j=1, 2, \cdots, n$，于是 $\mu I-A$ 有逆

$$\operatorname{adj}(\mu I-A)=\det(\mu I-A)(\mu I-A)^{-1},$$

记 $Z=[z_1z_2\cdots z_n]$ 是一个正交阵，$\Lambda=\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$，有

$$A=Z\Lambda Z^*,$$

故
$$\operatorname{adj}(\mu I-A)=\chi(\mu)Z(\mu-\Lambda)^{-1}Z^*,$$

或

$$\operatorname{adj}(\mu I-A)=Z\Delta(\mu)Z^*, \tag{13}$$

其中
$$\Delta(\mu)=\operatorname{diag}\left(\frac{\chi(\mu)}{\mu-\lambda_1}, \frac{\chi(\mu)}{\mu-\lambda_2}, \cdots, \frac{\chi(\mu)}{\mu-\lambda_n}\right),$$

等式(13)的两边都是 μ 的解析函数，因此对任何复数都成立。特别取 $\mu=\lambda_i$，

$$\Delta(\mu)=\Delta(\lambda_i)=\operatorname{diag}(0, 0, \cdots, \underset{\text{第 } i \text{ 列}}{\chi'(\lambda_i)}, 0, \cdots, 0),$$

于是
$$\operatorname{adj}(\lambda_i I-A)=Z\Delta(\lambda_i)Z^*=\chi'(\lambda_i)z_i z_i^*.$$ 证毕.

定理 2.8 将特征向量的分量与伴随阵元素联系起来。因此对于形式简单的矩阵，就有可能将特征向量的分量与矩阵的元素联系起来。1971 年 C. Paige 就给出了如下的定理[13].

定理 2.8 设对称三对角矩阵 T 的特征值为 $\theta_1, \theta_2, \cdots, \theta_n$，对应的标准化的正交特征向量为 s_1、s_2、$\cdots$、s_n，则对 $\mu\leqslant\nu$，$j=1$、2、$\cdots$、n 有

$$\chi'_{1,n}(\theta_j)s_{\mu j}s_{\nu j}=\chi_{1,\mu-1}(\theta_j)\beta_\mu\cdots\beta_{\nu-1}\chi_{\nu+1,n}(\theta_j), \tag{14}$$

这里 $s_{\mu j}$ 是 $\boldsymbol{s}_j$ 的第 μ 个分量. 特别当 T 不可约时(即 $\beta_i \neq 0$, $i=1、2、\cdots、n-1$)有

$$s_{\mu j}^2 = \chi_{1,\mu-1}(\theta_j)\chi_{\mu+1,n}(\theta_j)/\chi'_{1,n}(\theta_j). \tag{15}$$

证明 先证(14)式. 利用定理2.8有

$$\operatorname{adj}(\theta_j I - T) = \chi'_{1,n}(\theta_j)\boldsymbol{s}_j\boldsymbol{s}_j^*, \tag{16}$$

等式右边的第 μ 行第 ν 列即为 $\chi'_{1,n}(\theta_j)s_{\mu j}s_{\nu j}$; 等式左边的第 μ 行第 ν 列是 $\theta_j I - T$ 的第 ν 行第 μ 列的代数余子式, 因为 $\theta_j I - T$ 是对称的, 因此也是 $\theta_j I - T$ 的第 μ 行第 ν 列的代数余子式. 这个代数余子式记为 $\Delta_{\mu,\nu}$.

为了求 $\Delta_{\mu,\nu}$ 我们将 $\theta_j I - T$ 划去第 μ 行和第 ν 列后的矩阵表示如下:

$$\begin{pmatrix} \theta_j I - T_{1,\mu-1} & \begin{matrix} 0 \\ -\beta_{\mu-1} \end{matrix} & 0 & 0 \\ 0 & \begin{pmatrix} -\beta_\mu & & & \\ & -\beta_{\mu+1} & & \\ & & \ddots & \\ 0 & & & -\beta_{\nu-1} \end{pmatrix} & & 0 \\ 0 & & 0 & \begin{matrix} -\beta_\nu & 0 \\ \theta_j I - T_{\nu+1,n} \end{matrix} \end{pmatrix},$$

因此

$$\begin{aligned} \Delta_{\mu\nu} &= (-1)^{\mu+\nu}\det(\theta_j I - T_{1,\mu-1})(-1)^{\nu-\mu} \\ &\quad \times \beta_\mu\beta_{\mu+1}\cdots\beta_{\nu-1}\det(\theta_j I - T_{\nu+1,n}) \\ &= \chi_{1,\mu-1}(\theta_j)\beta_\mu\beta_{\mu+1}\cdots\beta_{\nu-1}\chi_{\nu+1,n}(\theta_j), \end{aligned}$$

(14)式获证.

当 $\nu=\mu$ 时, (16)式的右边第 μ 行第 μ 列为 $\chi'_{1,n}(\theta_j)s_{\mu j}^2$, 而左边的第 μ 行第 μ 列为 $\chi_{1,\mu-1}(\theta_j)\chi_{\mu+1,n}(\theta_j)$. 又因为 T 不可约时它没有重特征值, 因此 $\chi'_{1,n}(\theta_j) \neq 0$. 故(15)式成立. 证毕.

推论

$$s_{1j}s_{nj}\chi'(\theta_j)=\beta_1\beta_2\cdots\beta_{n-1}, \tag{17}$$

$$s_{1j}^2\chi'(\theta_j)=\chi_{2,n}(\theta_j), \tag{18}$$

$$s_{nj}^2\chi'(\theta_j)=\chi_{1,n-1}(\theta_j), \tag{19}$$

这里记 $\chi(\theta)=\chi_{1,n}(\theta)$.

证明 在定理中取 $\mu=1$, $\nu=n$ 即得(17), 取 $\mu=\nu=1$即得(18), 取 $\mu=\nu=n$ 得到(19). 证毕.

从(17)式可知, 如果 T 不可约, 那么 $\beta_1\beta_2\cdots\beta_{n-1}\neq 0$, $\chi'(\theta_j)\neq 0$, 因此 $s_{1j}s_{nj}\neq 0$, 即任意一个特征向量的第1个分量和最后一个分量不为0. 再由(18)、(19)两式知 $\chi_{2,n}(\theta_j)\neq 0$, $\chi_{1,n-1}(\theta_j)\neq 0$. 这也可从 Jacobi 矩阵根的分隔性质而得到.

现在来考虑矩阵的特征值反问题, 也即从已知 n 个特征值来确定一个 $n\times n$ 矩阵. 当然 $n\times n$ 矩阵有 n^2 个元素, 一般情况不能被 n 个特征值所确定. 因此还要提供另外的一些信息, 才能完全确定一个矩阵.

当 $n\times n$ 矩阵是一个对称三对角矩阵时, 要确定的矩阵元素只有 $2n-1$ 个, 因此需要附加的信息也较简单.

1967 年 Harry Hochstadt[14] 提出二种类型问题.

问题 1 给定二个序列 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 和 $\mu_1, \mu_2, \cdots, \mu_{n-1}$ 并且 $\lambda_i<\mu_i<\lambda_{i+1}$, $i=1, 2, \cdots, n-1$, 要求构造一个 $n\times n$ 对称三对角矩阵 $T_{1,n}$, $\beta_i>0$, 使得 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $T_{1,n}$ 的特征值, 而 $\mu_1, \mu_2, \cdots, \mu_{n-1}$ 是 $T_{2,n}$ 的特征值.

问题 2 给定 n 个由小到大的数 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 要求构造一个 $n\times n$ 对称三对角阵 T, 使得 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 T 的特征值, 但是 T 的元素满足:

$$\alpha_i=\alpha_{n+1-i},$$

$$\beta_i=\beta_{n-i}>0,$$

$$i=1, 2, \cdots, \left[\frac{n}{2}\right],$$

也即T是关于第二条主对角线(从右上到左下)也是对称的.

Hochstadt 证明问题1和问题2都至多有一个解. 1976年 Ole, H. Hald 证明问题1和问题2都至少有一个解,并且给出求解的方法[15],但是算法不稳定,1978年 C. de Boor 和 G. H. Golub,给出一种数值稳定的算法来求上述问题1和问题2的解[16],我们下面介绍的 B. N. Parlett 教授在[10]中给出的,利用 C. Paige 公式来构造T的方法. 在介绍这项方法之前,先来简单介绍一下对称三对角矩阵的特征值反问题实际背景.

考虑一个两端固定的弦的振动问题,它的数学模型可以化成下列边值问题:

$$\begin{cases} u''(x)-\sigma(x)u(x)=\lambda u(x); \\ u(0)=u(1)=0, \end{cases} \tag{20}$$

$\sigma(x)$是跟弦的密度有关的函数,如果弦是均匀的,那么$\sigma(x)=\sigma_0=\text{const}$. 当$\lambda$取固有值时,(20)有非零解$u(x)$,这是固有值对应弦的固有频率.

已知$\sigma(x)$,求固有值λ的问题,称为固有值问题或特征值问题. 反过来,如果通过实验手段测得弦的固有频率,计算得到固有值,从已知固有值来求$\sigma(x)$的问题,就称为特征值反问题.

用差分方法解边值问题(20),就把微分方程的特征值问题,化成矩阵的特征值问题,因此也把微分方程的特征值反问题,化成矩阵的特征值反问题. 例如$u''(x)$用中心差分$u(x_{i+1})-2u(x_i)+u(x_{i-1})/h^2$代替,这里$h=1/n$, $x_i=ih$,那么对应的矩阵是一个对称三角阵. 因此弦振动问题的特征值

反问题，就化成一个对称三对角矩阵的特征值反问题.

边值问题(20)是弦振动问题的数学模型，也可看作一个粒子的一维薛定谔方程，这时 $u(x)$ 表示粒子的波函数，$\sigma(x)$ 代表势函数，固有值 λ 代表在这个场中运动的粒子所具有的能级. 如果已知各个允许的能级，求势函数 $\sigma(x)$ 的问题也是对称三对角矩阵的特征值反问题.

上面介绍的简单的物理问题对应对称三对角矩阵的特征值反问题，可以设想，更复杂一些的某些物理问题，会对应更复杂一些矩阵的特征值反问题. 因此从七十年代以来矩阵特征值反问题的研究，被专家们所重视.

现在回到问题1、2的解上来. 设 T 是对称三对角矩阵，$\theta_1, \theta_2, \cdots, \theta_n$ 是它的特征值，$\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_n$ 是对应的特征向量，$S=[\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_n]$ 是正交阵，$\Lambda=\operatorname{diag}(\theta_1, \theta_2, \cdots, \theta_n)$，于是有

$$S^*TS=\Lambda,$$
$$T=S\Lambda S^*,$$

将 Λ 看作一个对称矩阵 A，利用§2的唯一归化定理知 T 和 S^* 的其他各列，完全被 Λ 和 S^* 的第1列 $(s_{11}, s_{12}, \cdots, s_{1n})^T$ 所确定，这里 s_{1i} 是 T 的第 i 个特征向量的第一个分量. 因此如果由 Λ 和反问题所给的附加信息能够确定 $(s_{11}, s_{12}, \cdots, s_{1n})$，那么就能完全确定 T.

对于问题1，已知 $\mu_1, \mu_2, \cdots, \mu_{n-1}$ 是 $T_{2,n}$ 的特征值，因此 $\chi_{2n}(\lambda)=\det(\lambda I-T_{2,n})=\prod\limits_{i=1}^{n-1}(\lambda-\mu_i)$，由(18)

$$s_{1j}^2\chi'(\theta_j)=\chi_{2n}(\theta_j),$$

即得
$$s_{1j}^2=\prod_{i=1}^{n-1}(\theta_j-\mu_i)\Big/\prod_{\substack{i=1\\i\neq j}}^{n}(\theta_j-\theta_i),$$

因为特征向量可以差一个符号．可取

$$s_{1j}=\sqrt{\prod_{i=1}^{n-1}(\theta_j-\mu_i)\Big/\prod_{\substack{i=1\\i\neq j}}^{n}(\theta_j-\theta_i)},$$

再由 Lanczos 过程(7)，就可以得到 T.

对于问题2，因为 T 是关于第2条主对角线对称的，

$$\alpha_i=\alpha_{n+1-i},\quad \beta_i=\beta_{n-i}>0,$$

于是若记矩阵 $\tilde{I}=(\boldsymbol{e}_n,\ \boldsymbol{e}_{n-1},\ \cdots,\ \boldsymbol{e}_2,\ \boldsymbol{e}_1)$，有

$$T=\tilde{I}T\tilde{I},$$

由

$$T\boldsymbol{s}_i=\lambda_i\boldsymbol{s}_i,$$

$$\tilde{I}T\tilde{I}\boldsymbol{s}_i=\lambda_i\boldsymbol{s}_i,$$

$$T\tilde{I}\boldsymbol{s}_i=\lambda_i\tilde{I}\boldsymbol{s}_i,$$

因为 $\beta_i>0$, T 不可约，T 的特征值都不相同，因此 $\tilde{I}\boldsymbol{s}_i$ 与 $\boldsymbol{s}_i$ 至多差一个符号，即

$$\tilde{I}\boldsymbol{s}_i=\varepsilon_i\boldsymbol{s}_i,\quad \varepsilon_i=\pm1,$$

由此可知

$$s_{ni}=\varepsilon_i s_{1i},$$

利用(17)

$$s_{1i}s_{ni}\chi'(\theta_i)=\beta_1\beta_2\cdots\beta_{n-1},$$

或

$$s_{1i}^2=\frac{\beta_1\beta_2\cdots\beta_{n-1}}{\varepsilon_i\chi'(\theta_i)},$$

但 $\chi'(\theta_i)=(\theta_i-\theta_1)\cdots(\theta_i-\theta_{i-1})(\theta_i-\theta_{i+1})\cdots(\theta_i-\theta_n)$,

$\operatorname{sign}\chi'(\theta_i)=(-1)^{n-i}$, $\beta_1\cdots\beta_{n-1}>0$, 故 $\varepsilon_i=(-1)^{n-i}$,

从而得到

$$s_{1i}=\sqrt{\beta_1\beta_2\cdots\beta_{n-1}\Big/\prod_{l=1}^{i-1}(\theta_i-\theta_l)\prod_{l=i+1}^{n}(\theta_l-\theta_i)},$$

由此也可求得 T.

现在我们把上面的理论，应用到如下的一个有实际意义的问题中去．考虑如图所示的 n 个质点，n 段弹簧的振动模

型，m_i 表示第 i 个质点的质量，k_i 表示第 i 段弹簧的弹性系数．若已知这个弹性系统的全部固有频率，又知道这个弹性系统的最后一段弹簧和最后一个质点去掉后的弹性系统的全部固有频率；同时知道 m_1，要求出 $k_1, k_2, \cdots, k_n$ 和 $m_2, m_3, \cdots, m_n$ 的值．

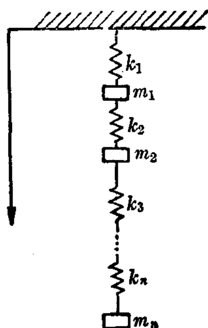

图 3

要解这个问题，先把振动方程列出来．设 l_i 表示第 i 段弹簧在静止状态时的长度，u_i 表示第 i 段弹簧的伸长，x_i 表示第 i 个质点的坐标．于是

$$x_i=\sum_{j=1}^{i} u_j+\sum_{j=1}^{i} l_j,$$

按照牛顿力学方程，对第 i 个质点成立

$$m_i\ddot{x}_i=k_{i+1}u_{i+1}-k_iu_i,$$

即

$$m_i\sum_{j=1}^{i}\ddot{u}_j=k_{i+1}u_{i+1}-k_iu_i,$$

或

$$\sum_{j=1}^{i}\ddot{u}_j=\frac{k_{i+1}}{m_i}u_{i+1}-\frac{k_i}{m_i}u_i,\quad i=1,\ 2,\ \cdots,\ n, \tag{21}$$

$$k_{n+1}=0,$$

从(21)可得

$$\ddot{u}_i=\frac{k_{i-1}}{m_{i-1}}u_{i-1}-\left(\frac{k_i}{m_{i-1}}+\frac{k_i}{m_i}\right)u_i$$

$$+\frac{k_{i+1}}{m_i}u_{i+1},\quad i=1,\ 2,\ \cdots,\ n, \tag{22}$$

其中规定 $\frac{k_0}{m_0}=\frac{k_1}{m_0}=0$.

若记 $a_i=-\left(\frac{k_i}{m_{i-1}}+\frac{k_i}{m_i}\right)$，$b_i=\frac{k_i}{m_i}$，$c_i=\frac{k_{i+1}}{m_i}$，

$$A_{1,n}=\begin{pmatrix} a_1 & c_1 & & & 0 \\ b_1 & a_2 & c_2 & & \\ & b_2 & \ddots & \ddots & \\ & & \ddots & \ddots & c_{n-1} \\ 0 & & & b_{n-1} & a_n \end{pmatrix},$$

$\boldsymbol{u}=(u_1, u_2, \cdots, u_n)^T$，则方程(22)可表示成

$$\ddot{\boldsymbol{u}}=A_{1,n}\boldsymbol{u},$$

若 $A_{1,n}$ 的特征值为 $-\lambda_l^2$, $l=1、2、\cdots、n$, 则 λ_l, $l=1、2、\cdots、n$, 即为这个系统的全部固有频率. 同时若

$$A_{1,n-1}=\begin{pmatrix} a_1 & c_1 & & & 0 \\ b_1 & a_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & c_{n-2} \\ 0 & & & b_{n-2} & a_{n-1} \end{pmatrix}$$

的特征值为 $-\mu_1^2, -\mu_2^2, \cdots, -\mu_{n-1}^2$, 则 $\mu_1, \mu_2, \cdots, \mu_{n-1}$ 即为去掉弹簧 k_n 和质量 m_n 后的系统的全部固有频率.

为了应用本节的理论，从 λ_l 和 μ_l 求 k_i、m_i, 我们先要将 $A_{1,n}$ 对称化. 若

$$D_n=\mathrm{diag}(1, \delta_1, \delta_2, \cdots, \delta_{n-1}),$$

有

$$D_n^{-1}A_{1,n-1}D_n$$

$$=\begin{pmatrix} a_1 & c_1\delta_1 & & & 0 \\ b_1/\delta_1 & a_2 & c_2\delta_2/\delta_1 & & \\ & b_2\delta_1/\delta_2 & \ddots & \ddots & \\ & & \ddots & \ddots & c_{n-1}\delta_{n-1}/\delta_{n-2} \\ 0 & & & b_{n-1}\delta_{n-2}/\delta_{n-1} & a_n \end{pmatrix},$$

如果取 δ_1、δ_2、…、δ_{n-1}，使得

$$b_i\delta_{i-1}/\delta_i = c_i\delta_i/\delta_{i-1}, \quad i=1, 2, \cdots, n-1,$$

就能使

$$B_{1,n} = D_n^{-1}A_{1,n}D_n$$

是一个对称三对角矩阵，此时

$$\delta_i = (b_1b_2\cdots b_i/c_1c_2\cdots c_i)^{\frac{1}{2}},$$

$$B_{1,n} = \begin{pmatrix} a_1 & \beta_1 & & & 0 \\ \beta_1 & a_2 & \beta_2 & & \\ & \beta_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1} \\ 0 & & & \beta_{n-1} & a_n \end{pmatrix},$$

其中 $\beta_i = \sqrt{b_ic_i} = \sqrt{k_ik_{i+1}}/m_i$.

同时可知 $B_{1,n-1}$ 相似于 $A_{1,n-1}$. 这样我们的问题化成已知 $B_{1,n}$ 和 $B_{1,n-1}$ 的特征值，以及 m_1 求 $m_2, \cdots, m_n$ 和 $k_1, \cdots, k_n$. 如果知道了 $B_{1,n}$ 的全部元素 $a_1, a_2, \cdots, a_n$ 和 $\beta_1, \beta_2, \cdots, \beta_{n-1}$，那么可以从如下递推关系式

$$\begin{aligned} &k_1 = -a_1m_1, \\ &k_i = \beta_{i-1}^2m_{i-1}^2/k_{i-1}, \\ &m_i = -k_i/(a_i + k_i/m_{i-1}), \quad i=2, 3, \cdots, n, \end{aligned} \tag{23}$$

求出 k_i 和 m_i.

由此问题化成已知 $B_{1,n}$ 和 $B_{1,n-1}$ 的特征值求 $B_{1,n}$ 的全部元素. 这就是 Hochstadt 的问题 1. 不过我们现在知道的是 $B_{1,n-1}$ 的特征值而不是 $B_{2,n}$ 的特征值.

如果 $s=(s_1, s_2, \cdots, s_n)$ 是 $B_{1,n}$ 特征向量构成的正交阵，

$$\Lambda = \mathrm{diag}(-\lambda_1^2, -\lambda_2^2, \cdots, -\lambda_n^2),$$

由

$$B_{1,n} = S\Lambda S^*,$$

可知 $B_{1,n}$ 完全被 S^* 的最后一列和 Λ 所完全确定. 但 S^* 的

最后一列为

$$\boldsymbol{g}_n=(s_{n1},\ s_{n2},\ \cdots,\ s_{nn})^T,$$

而利用公式(19),

$$s_{nj}^2\chi'(-\lambda_j^2)=\chi_{1,n-1}(-\lambda_j^2),$$

可以求出 $(s_{n1},\ s_{n2},\ \cdots,\ s_{nn})^T$, 这里 $\chi(\lambda)$ 和 $\chi_{1,n-1}(\lambda)$ 分别是 $B_{1,n}$ 和 $B_{1,n-1}$ 的特征多项式,即

$$s_{nj}^2=\chi_{1,n-1}(-\lambda_j^2)/\chi'(-\lambda_j^2)=\prod_{l=1}^{n-1}(\mu_l^2-\lambda_j^2)\Big/\prod_{\substack{l\neq j\\ l=1}}^{n}(\lambda_l^2-\lambda_j^2), \tag{24}$$

有了 $\boldsymbol{g}_n$ 后通过朝前推的 Lanczos 过程, 就可以求出 $B_{1,n}$ 的全部元素.

例: 对于 3 个质点和 3 个弹簧的系统,已知

$$\lambda_1=2,\quad \lambda_2=4,\quad \lambda_3=6,$$

$$\mu_1=3,\quad \mu_2=5,$$

求 k_1, k_2, k_3, m_2, m_3 用 m_1 的表示式.

解 $$s_{31}^2=\frac{(\mu_1^2-\lambda_1^2)(\mu_2^2-\lambda_1^2)}{(\lambda_2^2-\lambda_1^2)(\lambda_3^2-\lambda_1^2)}=0.2734375,$$

$$s_{32}^2=\frac{(\mu_1^2-\lambda_2^2)(\mu_2^2-\lambda_2^2)}{(\lambda_1^2-\lambda_2^2)(\lambda_3^2-\lambda_2^2)}=0.2625,$$

$$s_{31}^2=\frac{(\mu_1^2-\lambda_3^2)(\mu_2^2-\lambda_3^2)}{(\lambda_1^2-\lambda_3^2)(\lambda_2^2-\lambda_3^2)}=0.4640625,$$

得 $$\boldsymbol{g}_3=(0.5229125,\ 0.5123475,\ 0.6812213)^T,$$

$$\Lambda=\begin{pmatrix}-4 & 0 & 0\\ 0 & -16 & 0\\ 0 & 0 & -36\end{pmatrix},$$

$$\alpha_3=(\Lambda\boldsymbol{g}_3,\ \boldsymbol{g}_3)=-21.99999784,$$

$$\beta_2=\|\Lambda\boldsymbol{g}_3-\alpha_3\boldsymbol{g}_3\|=13.747726,$$

$$\boldsymbol{g}_2=(A\boldsymbol{g}_3-a_3\boldsymbol{g}_3)/\beta_2$$
$$=(0.6846531,\ 0.2236067,\ -0.6937220)^T,$$
$$a_2=(A\boldsymbol{g}_2,\ \boldsymbol{g}_2)=-20.0000,$$
$$\beta_1=\|A\boldsymbol{g}_2-a_2\boldsymbol{g}_2-\beta_2\boldsymbol{g}_3\|=7.4161985,$$
$$\boldsymbol{g}_1=(A\boldsymbol{g}_2-a_2\boldsymbol{g}_2-\beta_2\boldsymbol{g}_3)/\beta_1$$
$$=(0.5077524,\ -0.8291561,\ 0.2338539)^T,$$
$$a_1=(A\boldsymbol{g}_1,\ \boldsymbol{g}_1)=-14.0000,$$

得 $$B_{1,n}=\begin{pmatrix}-14.0000 & 7.4161985 & 0\\ 7.4161985 & -20.0000 & 13.747726\\ 0 & 13.747726 & -22.0000\end{pmatrix},$$

再用递推公式(23)得

$$k_1=-a_1m_1=14m_1,$$
$$k_2=\beta_1^2m_1^2/k_1=3.9285714m_1,$$
$$m_2=-k_2/(a_2+k_2/m_1)=0.24444444m_1,$$
$$k_3=\beta_2^2m_2^2/k_2=2.8746655m_1,$$
$$m_3=-k_3/(a_3+k_3/m_2)=0.280729m_1,$$

从公式(24)可以看出，已知的固有频率必须满足

$$\lambda_1<\mu_1<\lambda_2<\mu_2<\cdots<\mu_{n-1}<\lambda_n,$$

否则 s_{nj} 中会出现虚数，问题就没有解.

§5　解对称线性代数方程组的 Lanczos 算法

在第1章介绍的共轭斜量法，适用于系数矩阵是对称正定的或对称非负的方程组的解法，在实际问题中也要碰到系数矩阵对称不定的方程组，即系数矩阵有正的特征值，也有负的特征值，例如文献[17]中讨论的就是一种对称不定的方程组．对于对称不定的方程组，共轭斜量法失效了，首先因为

$(A\boldsymbol{x}, \boldsymbol{y})$此时不是 $\boldsymbol{x}, \boldsymbol{y}$ 的一种内积.

C. C. Paige and M. A. Saunders 在[18]和 B. N. Parlett 在[19]都介绍了用 Lanczos 过程来解这种方程组的方法，称为 Lanczos 算法. 对于方程组

$$A\boldsymbol{x}=\boldsymbol{b}, \tag{25}$$

A 是 $n\times n$ 实对称矩阵, Lanczos 算法的程式如下:

1. 取一个初始向量 $\boldsymbol{x}_0$, 计算 $\boldsymbol{r}_0=\boldsymbol{b}-A\boldsymbol{x}_0$;

2. 求$\|\boldsymbol{r}_0\|$和 $\boldsymbol{q}_1=\boldsymbol{r}_0/\|\boldsymbol{r}_0\|$;

3. 由 Lanczos 过程(7),

$$\beta_i\boldsymbol{q}_{i+1}=A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1},$$
$$i=1,\ 2,\ \cdots,\ j,$$
$$\beta_0\boldsymbol{q}_0=0,$$

求标准正交化序列 $\boldsymbol{q}_1,\ \boldsymbol{q}_2,\ \cdots,\ \boldsymbol{q}_{j+1},\ \alpha_1,\ \alpha_2,\ \cdots,\ \alpha_j,\ \beta_1,\ \beta_2,\ \cdots,\ \beta_j$, 其中

$$\alpha_i=(A\boldsymbol{q}_i,\ \boldsymbol{q}_i),$$
$$\beta_i=\|A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1}\|;$$

4. 由

$$T_j=\begin{pmatrix} \alpha_1 & \beta_1 & & 0 \\ \beta_1 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_{j-1} \\ 0 & & \beta_{j-1} & \alpha_j \end{pmatrix},$$

如果 $\det(T_j)\neq 0$, 求方程组

$$T_j\boldsymbol{f}_j=\|\boldsymbol{r}_0\|\boldsymbol{e}_1 \tag{26}$$

的解 $\boldsymbol{f}_j=(\delta_1^{(j)},\ \delta_2^{(j)},\ \cdots,\ \delta_j^{(j)})^T$;

5. 记矩阵 $Q_j=[\boldsymbol{q}_1,\ \boldsymbol{q}_2,\ \cdots,\ \boldsymbol{q}_j]$, 构造 $Q_j\boldsymbol{f}_j$, 则 $\boldsymbol{x}_j=Q_j\boldsymbol{f}_j+\boldsymbol{x}_0$ 是方程(25)的第 j 次近似解. 并成立误差估计式

$$\|A\boldsymbol{x}_j-\boldsymbol{b}\|=|\delta_j^{(j)}|\beta_j, \tag{27}$$

(27)式可以证明如下:过程(7)成立关系式

$$AQ_j - Q_jT_j = \beta_j \boldsymbol{q}_{j+1}\boldsymbol{e}_j^T,$$

于是

$$AQ_j\boldsymbol{f}_j - Q_jT_j\boldsymbol{f}_j = \beta_j\boldsymbol{q}_{j+1}\delta_j^{(j)},$$

$$A(\boldsymbol{x}_j - \boldsymbol{x}_0) - \|\boldsymbol{r}_0\|\boldsymbol{q}_1 = \beta_j\delta_j^{(j)}\boldsymbol{q}_{j+1},$$

$$A\boldsymbol{x}_j - \boldsymbol{b} = \beta_j\delta_j^{(j)}\boldsymbol{q}_{j+1},$$

故

$$\|A\boldsymbol{x}_j - \boldsymbol{b}\| = \beta_j|\delta_j^{(j)}|$$

成立.

上述 Lanczos 方法是将求方程组(25)的解的问题，化成求对称三对角方程组(26)的问题. 从理论上来说由于 $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\cdots$, $\boldsymbol{q}_n$ 的互相正交性知当 $j=n$ 时，$\beta_n=0$，即 $\boldsymbol{x}_n$ 为正确解；但在实际计算上，由于 $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\cdots$, $\boldsymbol{q}_n$ 的正交性消失，β_n 不一定为0,但很多计算表明,常常有某些 $j<n$,使 $\beta_j|\delta_j^{(j)}|$ 很小,参见[18], [19], [13]. 因此 Lanczos 方法求解对称方程组，仍是一种有效的方法，特别它可以充分利用矩阵 A 的稀疏性，用到 A 的时候只要求计算 $A\boldsymbol{v}$，即 A 与向量 $\boldsymbol{v}$ 的乘法，这一点不但可以节约存贮单元，而且可以容易被向量运算机来实现.

在计算过程中,如果预先知道 $\beta_j\delta_j^{(j)}$,那么只要等到 $\beta_j\delta_j^{(j)}$ 足够小时，才进行步骤4和5，就可大大节省计算量，因此 $\beta_j\delta_j^{(j)}$ 记为 Δ_j 称为这一算法的判据. 于是怎样不通过解方程组(26)，而直接计算 $\delta_j^{(j)}$ 的问题就提出来了. 文献[18]中也提出了一种直接计算 $\delta_j^{(j)}$ 的方法，在本章§3定理2.6给出一个更简单的求 $\delta_j^{(j)}$ 的公式,即为当 $\alpha=\|\boldsymbol{r}_0\|$ 时代入公式(11)得

$$\Delta_j = \beta_j\delta_j^{(j)} = (-1)^{j-1}\|\boldsymbol{r}_0\|\frac{\beta_1\beta_2\cdots\beta_j}{d_j},$$

并有递推公式

$$\Delta_j = -\Delta_{j-1}\frac{d_{j-1}}{d_j}\beta_j,$$

行列式 d_j 可以通过递推公式

$$d_i = \alpha_i d_{i-1} - \beta_{i-1}^2 d_{i-2}, \quad \beta_0 = 0$$

得到. 如果 $d_j=0$ 的话 d_{j+1} 不为 0, d_{j-1} 也不为 0, 可以计算 Δ_{j+1},

$$\Delta_{j+1} = \Delta_{j-1}\frac{d_{j-1}}{d_{j+1}}\beta_j\beta_{j+1} = -\Delta_{j-1}\frac{\beta_{j+1}}{\beta_j}.$$

对于 $\|A\boldsymbol{x}_j-\boldsymbol{b}\|$, 除了 (27) 所示的等式外, 还可以利用 $\beta_1\beta_2\cdots\beta_j$ 的极值性质(定理 2.5)得到下述估计

$$\|A\boldsymbol{x}_j-\boldsymbol{b}\| = \frac{\|\boldsymbol{r}_0\|}{|d_j|}\min_{\psi\in mP_j}\|\psi(T_k)\boldsymbol{e}_1\|,$$

k 是大于 j、小于等于 n 的任意正整数.

在[18]中指出, 当矩阵 A 是对称正定时, Lanczos 算法得到的 x_j 与共轭斜量法第 j 步近似 x_j 是相同的. 这一结论, 我们证明如下: 记方程(25)的正确解为 $\tilde{\boldsymbol{x}}$, Lanczos 算法得到的第 j 次近似解为 $\boldsymbol{x}_j$, 只要证明

$$(A(\tilde{\boldsymbol{x}}-\boldsymbol{x}_j),\ A^l\boldsymbol{r}_0) = 0, \quad l = 0, 1, 2, \cdots, j-1 \qquad (28)$$

就行了. 因 A 正定, 假定 Lanczos 算法前 j 步得到的 $\beta_1, \beta_2, \cdots, \beta_{j-1}$ 都不为 0, 此时 $Q_j=[\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j]$ 可以算得, $T_j=Q_j^TAQ_j$ 也是一个正定矩阵, 因此可以将 T_j 进行三角分解

$$T_j = L_jD_jL_j^T,$$

其中 L_j 是下三角阵, 对角元为 1, D_j 为对角阵. 记

$$P_j = Q_jL_j^{-T} = [\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_j],$$

$\boldsymbol{p}_i$ 是 $\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_i$ 的线性组合, 而 $\boldsymbol{q}_l$ 是 $\boldsymbol{r}_0, A\boldsymbol{r}_0, \cdots, A^{l-1}\boldsymbol{r}_0$ 的线性组合, 因此 $\boldsymbol{p}_i$ 是 $\boldsymbol{r}_0, A\boldsymbol{r}_0, \cdots, A^{i-1}\boldsymbol{r}_0$ 的线性组合. 反之由 $\beta_1, \beta_2, \cdots, \beta_{j-1}$ 不为 0, 可知 $A^{l-1}\boldsymbol{r}_0$ 是 $\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_l$ 的

线性组合，从而也是 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_l$ 的线性组合．这样(28)式等价于

$$P_j^T A(\tilde{\boldsymbol{x}} - \boldsymbol{x}_j) = 0, \tag{29}$$

因为
$$\boldsymbol{x}_j = Q_j \boldsymbol{f}_j + \boldsymbol{x}_0,$$

故
$$A(\tilde{\boldsymbol{x}} - \boldsymbol{x}_j) = \boldsymbol{r}_0 - AQ_j \boldsymbol{f}_j,$$

所以

$$\begin{aligned} P_j^T A(\tilde{\boldsymbol{x}} - \boldsymbol{x}_j) &= P_j^T \boldsymbol{r}_0 - P_j^T A Q_j \boldsymbol{f}_j \\ &= L_j^{-1} Q_j^T \boldsymbol{r}_0 - L_j^{-1} Q_j^T A Q_j \boldsymbol{f}_j \\ &= L_j^{-1} \|\boldsymbol{r}_0\| \boldsymbol{e}_1 - L_j^{-1} T_j \boldsymbol{f}_j = 0, \end{aligned}$$

这就证明了 $\boldsymbol{x}_j$ 也是共轭斜量法得到的第 j 步近似解．如果 $\beta_l = 0$，那么由(27)知 $\boldsymbol{x}_l$ 即为 $\tilde{\boldsymbol{x}}$，此时 $\boldsymbol{r}_0$、$A\boldsymbol{r}_0$、$\cdots$、$A^l \boldsymbol{r}_0$ 线性相关，从而共轭斜量法得到的第 l 步近似也是 $\tilde{\boldsymbol{x}}$．这就证明了，当 A 正定时，两种方法是等价的．

另外由
$$P_j^T A P_j = L_j^{-1} Q_j^T A Q_j L_j^{-T} = D_j,$$

也即 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_j$ 是 A-正交的向量组，容易知道共轭斜量法中得到的向量组 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$、$\cdots$、$\boldsymbol{q}_j$ 中的 $\boldsymbol{q}_i$ 与 $\boldsymbol{p}_i$ 至多差个常数倍．

在矩阵 A 对称不定的情况，Lanczos 算法解方程组，还有两个问题需要考虑．第一个问题是此时 T_j 不能保证可以进行三角分解，$T_j = L_j D_j L_j^T$，因为 T_j 不是正定，它的某些顺序主子式可能为 0，因此如何对方程组(26)求解就是一个问题了．当然可以采用选主元的消去法，不过那样一来，T_j 的三对角的形状会遭破坏，存贮量就要增加．

第二个问题是第 j 次近似 $\boldsymbol{x}_j$ 与第 $j+1$ 次近似 $\boldsymbol{x}_{j+1}$，不象共轭斜量法中那样有简单的表示式 $\boldsymbol{x}_{j+1} = \boldsymbol{x}_j + \alpha_j \boldsymbol{q}_j$，这是因为 $\boldsymbol{f}_{j+1}$ 的为首 j 个分量与 $\boldsymbol{f}_j$ 没有简单的关系，因此当已经

计算得 x_j 后，如果还要计算 x_{j+1}，就不能充分利用 x_j 提供的信息，这也是很大的浪费．怎样避免这种浪费，即充分利用已算得的 x_j，也是一个问题．

C. C. Paige 和 A. Saunders 在 [18] 中提出了一种称为 SYMMLQ 的方法，就是解决上述两个问题的一种方法．下面介绍 SYMMLQ 方法：

在 Lanczos 算法步骤 1、2、3 以后，对每个 j 得到对称三对角矩阵 T_j；将 T_j 进行 LQ 分解

$$T_j=\widetilde{L}_jP_j,$$

其中 L_j 是下三角矩阵，P_j 是正交矩阵．为了求 $\widetilde{L}_j$ 和 P_j，对 T_j 右乘一系列旋转阵 $H_{i,i+1}$，使得

$$T_jH_{12}H_{23}\cdots H_{j-1,j}=\widetilde{L}_j,$$

其中 $H_{i,i+1}=(h_{lk})$，除 $h_{ii}=c_i$，$h_{i+1,i+1}=-c_i$，$h_{i,i+1}=h_{i+1,i}=s_i$ 外其余元素都与单位阵的相同．右乘 $H_{i,i+1}$ 的目的是为了将 T_j 的第 1 条超对角线上第 i 个元素化成 0．我们以 5 阶矩阵为例来观察计算的程式．

$$T_5=\begin{pmatrix}\alpha_1 & \beta_1 & 0 & 0 & 0\\ \beta_1 & \alpha_2 & \beta_2 & 0 & 0\\ 0 & \beta_2 & \alpha_3 & \beta_3 & 0\\ 0 & 0 & \beta_3 & \alpha_4 & \beta_4\\ 0 & 0 & 0 & \beta_4 & \alpha_5\end{pmatrix},$$

$$T_5H_{12}=\begin{pmatrix}\alpha_1 & \beta_1 & 0 & 0 & 0\\ \beta_1 & \alpha_2 & \beta_2 & 0 & 0\\ 0 & \beta_2 & \alpha_3 & \beta_3 & 0\\ 0 & 0 & \beta_3 & \alpha_4 & \beta_4\\ 0 & 0 & 0 & \beta_4 & \alpha_5\end{pmatrix}\begin{pmatrix}c_1 & s_1 & 0 & 0 & 0\\ s_1 & -c_1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 1\end{pmatrix}$$

$$
=\begin{pmatrix}
\alpha_1c_1+\beta_1s_1 & \alpha_1s_1-\beta_1c_1 & 0 & 0 & 0\\
\beta_1c_1+\alpha_2s_1 & \beta_1s_1-\alpha_2c_1 & \beta_2 & 0 & 0\\
\beta_2s_1 & -\beta_2c_1 & \alpha_3 & \beta_3 & 0\\
0 & 0 & \beta_3 & \alpha_4 & \beta_4\\
0 & 0 & 0 & \beta_4 & \alpha_5
\end{pmatrix},
$$

于是取 $$s_1=\frac{\beta_1}{\sqrt{\alpha_1^2+\beta_1^2}},\quad c_1=\frac{\alpha_1}{\sqrt{\alpha_1^2+\beta_1^2}},$$

就能把 β_1 位置上元素化成 0，记

$$\gamma_1=\alpha_1c_1+\beta_1s_1,\quad \delta_2=\beta_1c_1+\alpha_2s_1,\quad \varepsilon_3=\beta_2s_1,$$

$$\tilde{\gamma}_2=\beta_1s_1-\alpha_2c_1,\quad \tilde{\delta}_3=-\beta_2c_1,$$

于是 $$T_5H_{12}=\begin{pmatrix}
\gamma_1 & & & & \\
\delta_2 & \tilde{\gamma}_2 & \beta_2 & & \\
\varepsilon_3 & \tilde{\delta}_3 & \alpha_3 & \beta_3 & \\
 & & \beta_3 & \alpha_4 & \beta_4\\
 & & & \beta_4 & \alpha_5
\end{pmatrix},$$

$$T_5H_{12}H_{23}=\begin{pmatrix}
\gamma_1 & & & & \\
\delta_2 & \tilde{\gamma}_2 & \beta_2 & & \\
\varepsilon_3 & \tilde{\delta}_3 & \alpha_3 & \beta_3 & \\
 & & \beta_3 & \alpha_4 & \beta_4\\
 & & & \beta_4 & \alpha_5
\end{pmatrix}\begin{pmatrix}
1 & & & & \\
 & c_2 & s_2 & & \\
 & s_2 & -c_2 & & \\
 & & & 1 & \\
 & & & & 1
\end{pmatrix}$$

$$=\begin{pmatrix}
\gamma_1 & 0 & 0 & 0 & 0\\
\delta_2 & \tilde{\gamma}_2c_2+\beta_2s_2 & \tilde{\gamma}_2s_2-\beta_2c_2 & 0 & 0\\
\varepsilon_3 & \tilde{\delta}_3c_2+\alpha_3s_2 & \tilde{\delta}_3s_2-\alpha_3c_2 & \beta_3 & 0\\
0 & \beta_3s_2 & -\beta_3c_2 & \alpha_4 & \beta_4\\
0 & 0 & 0 & \beta_4 & \alpha_5
\end{pmatrix},$$

取 $$s_2=\beta_2/(\tilde{\gamma}_2^2+\beta_2^2)^{\frac{1}{2}},\quad c_2=\tilde{\gamma}_2/(\tilde{\gamma}_2^2+\beta_2^2)^{\frac{1}{2}},$$

则 β_2 位置上的元素化成 0。记

$$\gamma_2=\tilde{\gamma}_2 c_2+\beta_2 s_2,$$
$$\delta_3=\tilde{\delta}_3 c_2+\alpha_3 s_2,$$
$$\varepsilon_4=\beta_3 s_2,$$
$$\tilde{\gamma}_3=\tilde{\delta}_3 s_2-\alpha_3 c_2,$$
$$\tilde{\delta}_4=-\beta_3 c_2.$$

从上述计算中可见：γ、δ、ε 在下次旋转时不改变，而 $\tilde{\gamma}$, $\tilde{\delta}$ 在下次旋转时要改变.

$$T_5 H_{12} H_{23}=\begin{pmatrix}\gamma_1 & 0 & 0 & 0 & 0\\ \delta_2 & \gamma_2 & 0 & 0 & 0\\ \varepsilon_3 & \delta_3 & \tilde{\gamma}_3 & \beta_3 & 0\\ 0 & \varepsilon_4 & \tilde{\delta}_4 & \alpha_4 & \beta_4\\ 0 & 0 & 0 & \beta_4 & \alpha_5\end{pmatrix},$$

$$T_5 H_{12} H_{23} H_{34}=\begin{pmatrix}\gamma_1 & 0 & 0 & 0 & 0\\ \delta_2 & \gamma_2 & 0 & 0 & 0\\ \varepsilon_3 & \delta_3 & \tilde{\gamma}_3 & \beta_3 & 0\\ 0 & \varepsilon_4 & \tilde{\delta}_4 & \alpha_4 & \beta_4\\ 0 & 0 & 0 & \beta_4 & \alpha_5\end{pmatrix}\begin{pmatrix}1 & & & & \\ & 1 & & & \\ & & c_3 & s_3 & \\ & & s_3 & -c_3 & \\ & & & & 1\end{pmatrix}$$

$$=\begin{pmatrix}\gamma_1 & 0 & 0 & 0 & 0\\ \delta_2 & \gamma_2 & 0 & 0 & 0\\ \varepsilon_3 & \delta_3 & \tilde{\gamma}_3 c_3+\beta_3 s_3 & \tilde{\gamma}_3 s_3-\beta_3 c_3 & 0\\ 0 & \varepsilon_4 & \tilde{\delta}_4 c_3+\alpha_4 s_3 & \tilde{\delta}_4 s_3-\alpha_4 c_3 & \beta_4\\ 0 & 0 & \beta_4 s_3 & -\beta_4 c_3 & \alpha_5\end{pmatrix},$$

要使(3, 4)位置上的元素为0，取

$$s_3=\beta_3/(\tilde{\gamma}_3^2+\beta_3^2)^{\frac{1}{2}},\quad c_3=\tilde{\gamma}_3/(\tilde{\gamma}_3^2+\beta_3^2)^{\frac{1}{2}},$$

记
$$\gamma_3=\tilde{\gamma}_3 c_3+\beta_3 s_3,$$
$$\delta_4=\tilde{\delta}_4 c_3+\alpha_4 s_3,$$
$$\varepsilon_5=\beta_4 s_3,$$

$$\tilde{\gamma}_4=\tilde{\delta}_4 s_3-\alpha_4 c_3,$$

$$\tilde{\delta}_5=-\beta_4 c_3,$$

$$T_5H_{12}H_{23}H_{34}=\begin{pmatrix}\gamma_1 & 0 & 0 & 0 & 0\\ \delta_2 & \gamma_2 & 0 & 0 & 0\\ \varepsilon_3 & \delta_3 & \gamma_3 & 0 & 0\\ 0 & \varepsilon_4 & \delta_4 & \tilde{\gamma}_4 & \beta_4\\ 0 & 0 & \varepsilon_5 & \tilde{\delta}_5 & \alpha_5\end{pmatrix},$$

总结上述过程，有一般的计算程式

$$\tilde{\gamma}_1=\alpha_1,\quad \tilde{\delta}_2=\beta_1,$$

则使用 $H_{k,k+1}$ 旋转时

$$s_k=\beta_k/(\tilde{\gamma}_k^2+\beta_k^2)^{1/2},\quad c_k=\tilde{\gamma}_k/(\tilde{\gamma}_k^2+\beta_k^2)^{1/2},$$

通过旋转得到

$$\left.\begin{aligned}\gamma_k&=\tilde{\gamma}_k c_k+\beta_k s_k=(\tilde{\gamma}_k^2+\beta_k^2)^{1/2},\\ \delta_{k+1}&=\tilde{\delta}_{k+1}c_k+\alpha_{k+1}s_k,\\ \varepsilon_{k+2}&=\beta_{k+1}s_k,\\ \tilde{\gamma}_{k+1}&=\tilde{\delta}_{k+1}s_k-\alpha_{k+1}c_k,\\ \tilde{\delta}_{k+2}&=-\beta_{k+1}c_k,\\ k&=1,\ 2,\ \cdots,\end{aligned}\right\}\tag{30}$$

按此公式计算

$$T_5H_{12}H_{23}H_{34}H_{45}=\begin{pmatrix}\gamma_1 & & & & \\ \delta_2 & \gamma_2 & & & \\ \varepsilon_3 & \delta_3 & \gamma_3 & & \\ 0 & \varepsilon_4 & \delta_4 & \gamma_4 & \\ 0 & 0 & \varepsilon_5 & \delta_5 & \tilde{\gamma}_5\end{pmatrix}=\tilde{L}_5,$$

而
$$P_5=H_{45}H_{34}H_{23}H_{12}.$$

如果 $\beta_k\neq 0$，则有 $\gamma_k>0$，但 $\tilde{\gamma}_{k+1}$ 可能为 0。

从上述推导还可知

$$T_6P_5^T=T_6H_{12}H_{23}H_{34}H_{45}=\begin{pmatrix}\gamma_1 &&&&& \\ \delta_2 & \gamma_2 &&&& \\ \varepsilon_3 & \delta_3 & \gamma_3 &&& \\ 0 & \varepsilon_4 & \delta_4 & \gamma_4 && \\ 0 & 0 & \varepsilon_5 & \delta_5 & \tilde{\gamma}_5 & \beta_5 \\ 0 & 0 & 0 & \varepsilon_6 & \tilde{\delta}_6 & \alpha_6\end{pmatrix},$$

其中 $\tilde{\delta}_6=-\beta_5c_4,\quad \varepsilon_6=\beta_5s_4,$

而 $P_6^T=P_5^TH_{56}.$

T_j 经过 LQ 分解后，$T_j=\tilde{L}_jP_j$，于是方程组(26)

$$T_j\boldsymbol{f}_j=\|\boldsymbol{r}_0\|\boldsymbol{e}_1$$

可化成 $\tilde{L}_jP_j\boldsymbol{f}_j=\|\boldsymbol{r}_0\|\boldsymbol{e}_1,$

令 $\tilde{\boldsymbol{z}}_j=P_j\boldsymbol{f}_j$，如果 $\tilde{\boldsymbol{\gamma}}_j\neq0$，即 $\det(\tilde{L}_j)\neq0$，那么可以由方程

$$\tilde{L}_j\tilde{\boldsymbol{z}}_j=\|\boldsymbol{r}_0\|\boldsymbol{e}_1 \tag{31}$$

唯一确定 $\tilde{\boldsymbol{z}}_j$. 如果知道了 β_j，由 $\gamma_j=(\tilde{\gamma}_j^2+\beta_j^2)^{1/2}$，可构造一个新的下三角阵

$$L_j=\begin{pmatrix}\gamma_1 &&&&& \\ \delta_2 & \gamma_2 &&&& \\ \varepsilon_3 & \delta_3 & \ddots &&& \\ & \varepsilon_4 & \ddots & \ddots && \\ && \ddots & \ddots & \gamma_{j-1} & \\ &&& \varepsilon_j & \delta_j & \gamma_j\end{pmatrix},$$

与 $\tilde{L}_j$ 只差一个对角元，当 $\beta_j\neq0$ 时 $\det(L_j)\neq0$，如果令 $\boldsymbol{z}_j=(\zeta_1,\ \zeta_2,\ \cdots,\ \zeta_j)^T$，$\boldsymbol{z}_j$ 可由

$$L_j\boldsymbol{z}_j=\|\boldsymbol{r}_0\|\boldsymbol{e}_1$$

唯一确定．因为 L_j 是下三角阵，因此由

$$L_{j+1}\boldsymbol{z}_{j+1}=\|\boldsymbol{r}_0\|\boldsymbol{e}_1$$

确定的 $\boldsymbol{z}_{j+1}$ 必有 $\boldsymbol{z}_{j+1}=(\boldsymbol{z}_j^T,\ \zeta_{j+1})^T$，即 $\boldsymbol{z}_{j+1}$ 的为首 j 个分量

为 $\boldsymbol{z}_j$. 同时由方程(31), 可知

$$\tilde{\boldsymbol{z}}_j=(\zeta_1,\ \zeta_2,\ \cdots,\ \zeta_{j-1},\ \tilde{\zeta}_j)^T,$$

并且当 $\tilde{\gamma}_j\neq 0$ 时有

$$\tilde{\gamma}_j\tilde{\zeta}_j=\gamma_j\zeta_j.$$

再由 $\boldsymbol{x}_j=Q_j\boldsymbol{f}_j+\boldsymbol{x}_0=Q_jP_j^T\tilde{\boldsymbol{z}}_j+\boldsymbol{x}_0,$

需要分析 $Q_jP_j^T$. 由

$$Q_{j+1}P_{j+1}^T=(Q_j,\ q_{j+1})\begin{pmatrix}P_j^T & 0\\ 0 & 1\end{pmatrix}H_{j,j+1}$$

$$=(Q_jP_j^T,\ \boldsymbol{q}_{j+1})\begin{pmatrix}I & & 0\\ & c_j & s_j\\ 0 & s_j & -c_j\end{pmatrix}, \tag{32}$$

可知 $Q_{j+1}P_{j+1}^T$ 的为首 $j-1$ 列与 $Q_jP_j^T$ 为首 $j-1$ 列是完全相同的. 记

$$Q_jP_j^T=(\boldsymbol{w}_1,\ \boldsymbol{w}_2,\ \cdots,\ \boldsymbol{w}_{j-1},\ \tilde{\boldsymbol{w}}_j),$$

于是 $Q_{j+1}P_{j+1}^T=(\boldsymbol{w}_1,\ \boldsymbol{w}_2,\ \cdots,\ \boldsymbol{w}_{j-1},\ \boldsymbol{w}_j,\ \tilde{\boldsymbol{w}}_{j+1}),$

比较(32)的两边有

$$\left.\begin{aligned}\boldsymbol{w}_j&=c_j\tilde{\boldsymbol{w}}_j+s_j\boldsymbol{q}_{j+1},\\ \tilde{\boldsymbol{w}}_{j+1}&=s_j\tilde{\boldsymbol{w}}_j-c_j\boldsymbol{q}_{j+1},\end{aligned}\right\} \tag{33}$$

因为 $P_1^T=1$, 因此 $\tilde{\boldsymbol{w}}_1=\boldsymbol{q}_1$.

如记 $x_j^L=(\boldsymbol{w}_1,\ \boldsymbol{w}_2,\ \cdots,\ \boldsymbol{w}_j)\boldsymbol{z}_j+\boldsymbol{x}_0,$

就有

$$\boldsymbol{x}_j^L=\boldsymbol{x}_{j-1}^L+\zeta_j\boldsymbol{w}_j, \tag{34}$$

而 Lanczos 算法的近似解

$$\boldsymbol{x}_j=(\boldsymbol{w}_1,\ \boldsymbol{w}_2,\ \cdots,\ \boldsymbol{w}_{j-1},\ \tilde{\boldsymbol{w}}_j)\tilde{\boldsymbol{z}}_j+\boldsymbol{x}_0=\boldsymbol{x}_{j-1}^L+\tilde{\zeta}_j\tilde{\boldsymbol{w}}_j. \tag{35}$$

这就是 SYMMLQ 方法求 $\boldsymbol{x}_j$ 的过程, 在实际计算中, $\boldsymbol{x}_j^L$ 也是一个近似解, 在迭代过程中, 可以在前面只用(34),

求x_1^L、x_2^L、…、x_{j-1}^L，最后一步才用(35)求x_j，归纳一下SYMMLQ方法的步骤如下：

1．取初始近似向量x_0，求$r_0=b-Ax_0$，

计算$\|r_0\|$，$q_1=r_0/\|r_0\|$，$\tilde{w}_1=q_1$，

$\alpha_1=(Aq_1, q_1)$，$\tilde{\gamma}_1=\alpha_1$，求$Aq_1-\alpha_1 q_1$，

计算$\|Aq_1-\alpha_1 q_1\|=\beta_1$，

如果$\beta_1=0$，则$x_0+\frac{\|r_0\|}{\alpha_1}q_1\to x$，转5，

否则计算$q_2=(Aq_1-\alpha_1 q_1)/\beta_1$，$(Aq_2, q_2)=\alpha_2$，

$d_1=\alpha_1$，如果$d_1\neq 0$，则计算$\Delta_1=\frac{\|r_0\|\beta_1}{|d_1|}$，

$d_2=\alpha_1\alpha_2-\beta_1^2$，$\varepsilon_1=\varepsilon_2=\delta_1=0$，$\tilde{\delta}_2=\beta_1$，

$\gamma_1=(\tilde{\gamma}_1^2+\beta_1^2)^{1/2}$，$\zeta_1=\|r_0\|/\gamma_1$，

$s_1=\beta_1/\gamma_1$，$c_1=\tilde{\gamma}_1/\gamma_1$，

$\tilde{\gamma}_2=\tilde{\delta}_2 s_1-\alpha_2 c_1$，

$w_1=c_1\tilde{w}_1+s_1 q_2$，$\tilde{w}_2=s_1\tilde{w}_1-c_1 q_2$，

$x_1^L=x_0+\zeta_1 w_1$，

求$Aq_2-\alpha_2 q_2-\beta_1 q_1$，

计算$\beta_2=\|Aq_2-\alpha_2 q_2-\beta_1 q_1\|$，

如果$d_2\neq 0$及$\beta_2=0$，则$\tilde{\zeta}_2=-\delta_2\zeta_1/\tilde{\gamma}_2$，

$x_2=x_1^L+\tilde{\zeta}_2\tilde{w}_2\to x$转5.

如果$\beta_2\neq 0$，计算$\Delta_2=\frac{\|r_0\|\beta_1\beta_2}{|d_2|}$

$2\to i$转2.

如果$d_2=0$，$2\to i$转2.

2．$\varepsilon_{i+1}=\beta_i s_{i-1}$，$\tilde{\delta}_{i+1}=-\beta_i c_{i-1}$，

$\gamma_i=(\tilde{\gamma}_i^2+\beta_i^2)^{1/2}$，$s_i=\beta_i/\gamma_i$，$c_i=\tilde{\gamma}_i/\gamma_i$，

$q_{i+1}=(Aq_i-\alpha_i q_i-\beta_{i-1}q_{i-1})/\beta_i$，

$\alpha_{i+1}=(A\boldsymbol{q}_{i+1},\ \boldsymbol{q}_{i+1})$,

$\delta_{i+1}=\tilde{\delta}_{i+1}c_i+\alpha_{i+1}s_i$,

$\tilde{\gamma}_{i+1}=\tilde{\delta}_{i+1}s_i-\alpha_{i+1}c_i$,

$\boldsymbol{w}_i=\tilde{\boldsymbol{w}}_i c_i+\boldsymbol{q}_{i+1}s_i$,

$\tilde{\boldsymbol{w}}_{i+1}=\tilde{\boldsymbol{w}}_i s_i-\boldsymbol{q}_{i+1}c_i$,

$$\zeta_i=\frac{-\varepsilon_i\zeta_{i-2}-\delta_i\zeta_{i-1}}{\gamma_i},$$

$\boldsymbol{x}_i^L=\boldsymbol{x}_{i-1}^L+\zeta_i\boldsymbol{w}_i$.

3．构造 $\boldsymbol{g}_{i+1}=A\boldsymbol{q}_{i+1}-\alpha_{i+1}\boldsymbol{q}_{i+1}-\beta_i\boldsymbol{q}_i$,

计算 $\beta_{i+1}=\|\boldsymbol{g}_{i+1}\|$,

计算 $d_{i+1}=\alpha_{i+1}d_i-\beta_i^2 d_{i-1}$,

如果 $d_{i+1}=0$, 则 $i+1\to i$ 转 2,

如果 $d_{i+1}\neq 0$, 计算

$$\Delta_{i+1}=\begin{cases}\Delta_i\beta_{i+1}\left|\dfrac{d_i}{d_{i+1}}\right|, & \text{当 } d_i\neq 0,\\[2ex] \Delta_{i-1}\beta_i\beta_{i+1}\left|\dfrac{d_{i-1}}{d_{i+1}}\right|, & \text{当 } d_i=0,\end{cases}$$

如果 $\Delta_{i+1}<\varepsilon$ 转 4, 否则 $i+1\to i$ 转 2.

4．计算

$\tilde{\zeta}_{i+1}=-(\varepsilon_{i+1}\zeta_{i-1}+\delta_{i+1}\zeta_i)/\tilde{\gamma}_{i+1}$,

$\boldsymbol{x}_{i+1}=\boldsymbol{x}_i^L+\tilde{\zeta}_{i+1}\tilde{\boldsymbol{w}}_{i+1}\to\boldsymbol{x}$.

5．输出计算结果 $\boldsymbol{x}$; 停止.

第 3 章

解特征值问题的 QL 方法

计算中小型对称矩阵的特征值的一个比较有效的方法，是先将这个矩阵，通过第 2 章 § 2 所介绍的正交相似变换化成一个对称三对角矩阵，然后再用带特别选取的位移的 QL 方法(或 QR 方法)求得这个对称三对角矩阵的特征值.

熟知的 Givens 方法，是将矩阵化为对称三对角矩阵后，再用对分法求后者的特征值，但是对分法的收敛速度是一次的，而带特别选取的位移的 QL 方法(或 QR 方法)，收敛速度一般是三次的，并且每次计算量也是相当于对分法，因此这种 QL 方法是更有效.

本章要介绍 QL 方法有关的理论分析，指出它的收敛速度为何一般是三次.

§ 1　QL 方法的一般性质

设 A 是 $n\times n$ 实对称矩阵，σ 是实参数，将 $A-\sigma I$ 分解成

$$A-\sigma I=QL,$$

其中 Q 是 $n\times n$ 正交矩阵，L 是下三角阵. 记

$$\hat{A}=LQ+\sigma I,$$

于是 $$Q\hat{A}Q^*=QL+\sigma I=A-\sigma I+\sigma I=A,$$

故

$$\hat{A}=Q^*AQ, \tag{1}$$

可知 $\hat{A}$ 仍为实对称矩阵，并且正交相似于 A.

给定实数 σ_k, $k=1, 2, \cdots$, 令 $A_1=A$,

$$A_k-\sigma_k I=Q_k L_k, \tag{2}$$

Q_k 是正交阵 L_k 是下三角阵,

$$A_{k+1}=L_k Q_k+\sigma_k I \tag{3}$$

形成了矩阵序列$\{A_k\}$, 如果 A_k 趋于一个对角阵, 那么 A 的全部特征值就得到了. 这样求 A 的特征值的方法, 就称为带位移$\{\sigma_k\}$的 QL 方法. 如果在(2)中的分解是

$$A_k-\sigma_k I=Q_k R_k,$$

Q_k 是正交阵, R_k 是上三角阵, 且

$$A_{k+1}=R_k Q_k+\sigma_k I,$$

那么这样的求特征值的方法, 称为带位移 $\{\sigma_k\}$ 的 QR 方法.

现在来看 QL 分解是怎么进行的, 记 $Q=[\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_n]$, $A-\sigma I=[\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n]$,

$$L=\begin{pmatrix} l_{11} & & & & \\ l_{21} & l_{22} & & \text{\Large 0} & \\ l_{31} & l_{32} & l_{33} & & \\ \cdots & \cdots & \cdots & \cdots & \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{nn} \end{pmatrix},$$

比较 $A-\sigma I=QL$ 两边知

$$\boldsymbol{a}_n=l_{nn}\boldsymbol{q}_n,$$

因 $\|\boldsymbol{q}_n\|=1$, 故 $l_{nn}=\|\boldsymbol{a}_n\|$, 如果取 $l_{nn}>0$, 那么由 $\boldsymbol{a}_n$ 唯一确定 l_{nn} 和 $\boldsymbol{q}_n$. 再由

$$\boldsymbol{a}_{n-1}=\boldsymbol{q}_n l_{n,n-1}+\boldsymbol{q}_{n-1}l_{n-1,n-1},$$

或
$$\boldsymbol{q}_{n-1}l_{n-1,n-1}=\boldsymbol{a}_{n-1}-\boldsymbol{q}_n l_{n,n-1},$$

根据$(\boldsymbol{q}_{n-1}, \boldsymbol{q}_n)=0$, 得

$$l_{n,n-1}=(\boldsymbol{a}_{n-1}, \boldsymbol{q}_n),$$

同样确定
$$l_{n-1,n-1}=\|\boldsymbol{a}_{n-1}-\boldsymbol{q}_n l_{n,n-1}\|,$$

$$\boldsymbol{q}_{n-1}=(\boldsymbol{a}_{n-1}-\boldsymbol{q}_n l_{n,n-1})/l_{n-1,n-1},$$

一般地有

$$\boldsymbol{q}_k l_{kk}=\boldsymbol{a}_k-\boldsymbol{q}_{k+1}l_{k+1,k}-\boldsymbol{q}_{k+2}l_{k+2,k}-\cdots-\boldsymbol{q}_n l_{n,k},$$

其中

$$l_{jk}=(\boldsymbol{a}_k,\ \boldsymbol{q}_j),$$

$$j=k+1,\ k+2,\ \cdots,\ n,$$

$$l_{k,k}=\left\|\boldsymbol{a}_k-\sum_{j=k+1}^{n}\boldsymbol{q}_j l_{jk}\right\|,$$

$$\boldsymbol{q}_k=\left(\boldsymbol{a}_k-\sum_{j=k+1}^{n}\boldsymbol{q}_j l_{jk}\right)\Big/l_{k,k},$$

$$k=n-1,\ n-2,\ \cdots,\ 1.$$

如果 $l_{k,k}=0$，那么上述过程就要中断．这时

$$\boldsymbol{a}_k-\sum_{j=k+1}^{n}\boldsymbol{q}_j l_{jk}=0,$$

也即 $\boldsymbol{a}_k$、$\boldsymbol{a}_{k+1}$、…、$\boldsymbol{a}_n$ 线性相关，于是 $\det(A-\sigma I)=0$，说明 σ 是 A 的一个特征值．

反之如果 σ 是 A 的一个特征值，则至少存在一个 k，使 $l_{k,k}=0$．实际上若 σ 是 A 的特征值，那么 $\det(A-\sigma I)=0$，从而至少有一个 k，使得 $\boldsymbol{a}_k$、$\boldsymbol{a}_{k+1}$、…、$\boldsymbol{a}_n$ 线性相关．但

$$l_{kk}\boldsymbol{q}_k=\boldsymbol{a}_k-\sum_{j=k+1}^{n}\boldsymbol{q}_j l_{jk},$$

有

$$\left(\boldsymbol{a}_k-\sum_{j=k+1}^{n}\boldsymbol{q}_j l_{jk},\ \boldsymbol{q}_i\right)=0,\quad i=k+1,\ \cdots,\ n,$$

因此 $\sum_{j=k+1}^{n}\boldsymbol{q}_j l_{jk}$ 是 $\{\boldsymbol{a}_{k+1},\ \boldsymbol{a}_{k+2},\ \cdots,\ \boldsymbol{a}_n\}$ 中对 $\boldsymbol{a}_k$ 的最佳逼近，因此 $\boldsymbol{a}_k-\sum_{j=k+1}^{n}\boldsymbol{q}_j l_{jk}=0$，从而推出

$$l_{k,k}=0,$$

于是关于 QL 方法有

定理 3.1 存在一个正整数 k，使 $l_{kk}=0$ 的充分必要条件

是：σ 为 A 的特征值.

如果 σ 不是 A 的特征值，那么 l_{11}, l_{22}, $\cdots$, l_{nn} 都不为 0，这时 $A-\sigma I$ 的 QL 分解中，求 $\boldsymbol{q}_n$, $\boldsymbol{q}_{n-1}$, $\cdots$, $\boldsymbol{q}_1$ 的过程，即为将 $\boldsymbol{a}_n$, $\boldsymbol{a}_{n-1}$, $\cdots$, $\boldsymbol{a}_1$ 正交化的过程. 并且 Q 和 L 是唯一确定的.

在计算中如果遇到 $l_{kk}=0$，但计算还要进行下去，那么可取 $\boldsymbol{q}_k$ 为任意与 $\boldsymbol{q}_{k+1}$, $\boldsymbol{q}_{k+2}$, $\cdots$, $\boldsymbol{q}_n$ 正交的单位向量. 然后继续求 $\boldsymbol{q}_{k-1}$, $\cdots$, $\boldsymbol{q}_1$.

QL 方法的第二个重要性质是：

定理 3.2 如果 A 是带形矩阵，则 $\hat{A}$ 也是同样带形的矩阵，即保持原来的带宽.

证明 由

$$\boldsymbol{q}_n=\frac{1}{l_{nn}}\boldsymbol{a}_n,$$

$$\boldsymbol{q}_{n-1}=\frac{1}{l_{n-1,n-1}}(\boldsymbol{a}_{n-1}-l_{n,n-1}\boldsymbol{q}_n),$$

$$\cdots\cdots\cdots\cdots$$

$$\boldsymbol{q}_k=\frac{1}{l_{kk}}\left(\boldsymbol{a}_k-\sum_{j=k+1}^{n}\boldsymbol{q}_j l_{jk}\right),$$

可知 $Q=[\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_n]$ 的对角线以上部分是保持与 A 相同的带宽. 因为 L 是下三角阵，因此 LQ 在对角线以上部分仍保持与 A 相同的带宽. 但 $\hat{A}=LQ+\sigma I$ 是一个对称矩阵，因此它的对角线以下部分，也保持与 A 相同的带宽. 证毕.

推论 若 A 是对称三对角矩阵，则 $\hat{A}$ 也是对称三对角矩阵.

下面考察 QL 方法跟 QR 方法的关系. 对于矩阵 $A-\sigma I$ 进行 QR 分解，即

$$A-\sigma I=QR,$$

Q 是正交阵, R 是上三角阵. 对同一个矩阵 $A-\sigma I$, 进行 QR 分解得到的正交阵 Q, 与进行 QL 分解得到的正交阵显然一般是不相同的. 因此我们为了区别这两个正交阵, 我们用 $Q_R(B)$ 表示对 B 进行 QR 分解得到的正交阵, R_B 表此分解的上三角阵, 而用 $Q_L(B)$ 表示对 B 进行 QL 分解而得到的正交阵, L_B 表此分解的下三角阵.

记 $\tilde{I}=(\boldsymbol{e}_n, \boldsymbol{e}_{n-1}, \cdots, \boldsymbol{e}_1)$, 有 $\tilde{I}^{-1}=\tilde{I}$.

引理 3.1 设 $\det(B)\neq 0$, $\tilde{B}=\tilde{I}B\tilde{I}$, 则

$$Q_L(\tilde{B})=\tilde{I}Q_R(B)\tilde{I}, \quad L_{\tilde{B}}=\tilde{I}R_B\tilde{I}.$$

证明 $\tilde{B}=\tilde{I}B\tilde{I}=\tilde{I}Q_R(B)R_B\tilde{I}=\tilde{I}Q_R(B)\tilde{I}\tilde{I}R_B\tilde{I}$,
因为 $\tilde{I}R_B\tilde{I}$ 是一个下三角阵, 而 $\tilde{I}Q_R(B)\tilde{I}$ 是一个正交阵, 故上式表示了 $\tilde{B}$ 的 QL 分解, 但因 $\det(\tilde{B})=\det(B)\neq 0$, 因此 QL 分解是唯一的, 故

$$Q_L(\tilde{B})=\tilde{I}Q_R(B)\tilde{I}.$$

同时还知道 $L_{\tilde{B}}=\tilde{I}R_B\tilde{I}$. 证毕.

定理 3.3 设矩阵序列 $\{A_k\}$, $\{\tilde{A}_k\}$ 分别是使用 QL 方法和 QR 方法所得的序列, 所带的位移都是 $\{\sigma_k\}$. 如果 $\tilde{A}_1=\tilde{I}A_1\tilde{I}$, 则

$$\tilde{A}_k=\tilde{I}A_k\tilde{I}$$

对所有 $k>1$ 成立.

证明

$$\begin{aligned}
&A_1-\sigma_1 I=Q_L(A_1)L_{A_1},\\
&\tilde{A}_1-\sigma_1 I=Q_R(\tilde{A}_1)R_{\tilde{A}_1},\\
&A_2=L_{A_1}Q_L(A_1)+\sigma_1 I,\\
&\tilde{I}A_2\tilde{I}=\tilde{I}L_{A_1}\tilde{I}\tilde{I}Q_L(A_1)\tilde{I}+\sigma_1 I\\
&\qquad=R_{\tilde{A}_1}Q_R(\tilde{A}_1)+\sigma_1 I=\tilde{A}_2,
\end{aligned}$$

同理可证 $k=3, 4, \cdots$ 有

$$\tilde{A}_k=\tilde{I}A_k\tilde{I}. \text{ 证毕.}$$

从定理 3.3 可知，对某个矩阵进行 QL 方法求特征值与对这个矩阵进行$\begin{pmatrix}1 & 2 & 3 & \cdots & n\\ n & n-1 & n-2 & \cdots & 1\end{pmatrix}$置换相似变换后的矩阵进行 QR 方法是一样. 在历史上是先提出 QR 方法的，但是对于有些矩阵它的元素的数量级小的在左上角，逐渐向右下角增大，对于这种矩阵，计算经验表明，使用 QL 方法，常更加有利. 这种矩阵在工程技术问题中，常常遇到，因此 QL 方法有独立讨论的必要.

现来看 QL 方法与乘幂法、反迭代法的关系.

对于带位移 $\{\sigma_k\}$ 的 QL 方法:

$$A_1=A,$$

$$A_{k+1}=Q_k^*A_kQ_k=Q_k^*Q_{k-1}^*\cdots Q_1^*A_1Q_1Q_2\cdots Q_k,$$

记
$$P_k=Q_1Q_2\cdots Q_k,$$

故
$$A_{k+1}=P_k^*A_1P_k.$$

乘幂法：取单位向量 $\boldsymbol{v}_1$ 作为初始向量,

$$\boldsymbol{v}_{k+1}=(A-\sigma_k)\boldsymbol{v}_k/\nu_k, \tag{4}$$

取 ν_k 使得 $\|\boldsymbol{v}_{k+1}\|=1$.

反迭代法：取单位向量 $\boldsymbol{u}_1$ 作为初始向量,

$$(A-\sigma_k)\boldsymbol{u}_{k+1}=\boldsymbol{u}_k\tau_k, \tag{5}$$

取 τ_k 使 $\|\boldsymbol{u}_{k+1}\|=1$.

定理 3.4 如果序列$\{\sigma_k\}$中没有一个 σ_k 是 A 的特征值，并且若 $\boldsymbol{v}_1=\boldsymbol{e}_n$, $\boldsymbol{u}_1=\boldsymbol{e}_1$, 则

$$\boldsymbol{v}_{k+1}=P_k\boldsymbol{e}_n,\quad \boldsymbol{u}_{k+1}=P_k\boldsymbol{e}_1.$$

证明 因为 $A_{k+1}=P_k^*AP_k$,

故
$$P_kA_{k+1}=AP_k,$$

$$\begin{aligned}P_{k+1}L_{k+1}&=P_kQ_{k+1}L_{k+1}\\&=P_k(A_{k+1}-\sigma_{k+1}I)=(A-\sigma_{k+1}I)P_k,\end{aligned}$$

$$P_{k+1}L_{k+1}\boldsymbol{e}_n=(A-\sigma_{k+1}I)P_k\boldsymbol{e}_n,$$

$$l_{nn}^{(k+1)}P_{k+1}\boldsymbol{e}_n=(A-\sigma_{k+1}I)P_k\boldsymbol{e}_n,$$

因为 $\|P_{k+1}\boldsymbol{e}_n\|=1$，为了与(4)相比较，将上式写成

$$P_{k+1}\boldsymbol{e}_n=(A-\sigma_{k+1}I)P_k\boldsymbol{e}_n/l_{nn}^{(k+1)}, \tag{6}$$

又由

$$P_1\boldsymbol{e}_n=Q_1\boldsymbol{e}_n=(A-\sigma_1I)\boldsymbol{e}_n/l_{nn}^{(1)}=(A_1-\sigma_1I)\boldsymbol{v}_1/l_{nn}^{(1)},$$

因为 $\|P_1\boldsymbol{e}_n\|=1$，故 $\boldsymbol{v}_2=P_1\boldsymbol{e}_n$，再从(6)就知道一般

$$\boldsymbol{v}_{k+1}=P_k\boldsymbol{e}_n,$$

且知 $l_{nn}^{(k)}=\nu_k$.

另外 $$L_{k+1}^*P_{k+1}^*=P_k^*(A-\sigma_{k+1}I),$$

或 $$P_kL_{k+1}^*=(A-\sigma_{k+1}I)P_{k+1},$$

$$P_kL_{k+1}^*\boldsymbol{e}_1=(A-\sigma_{k+1}I)P_{k+1}\boldsymbol{e}_1,$$

$$l_{11}^{(k+1)}P_k\boldsymbol{e}_1=(A-\sigma_{k+1}I)P_{k+1}\boldsymbol{e}_1,$$

易知 $$\boldsymbol{u}_{k+1}=P_k\boldsymbol{e}_1,\ l_{11}^{(k)}=\tau_k.$$ 证毕.

从这个定理可知当 $\sigma_k=\sigma=\text{const}$ 时，P_k 的第一列常收敛到使 $1/|\lambda_i-\sigma|$ 达到最大的那个特征值 λ_i 所对应的特征向量，而且从

$$A_{k+1}=P_k^*AP_k,$$

可知 A_{k+1} 的元素 $a_{11}^{(k+1)}\to\lambda_i$, $a_{l1}^{(k+1)}=a_{1l}^{(k+1)}\to0$, $l=2,3,\cdots,n$.

同样可知 P_k 的第 n 列，当 $\sigma_k=\sigma=\text{const}$ 时，收敛到使 $|\lambda_i-\sigma|$ 最大的特征值 λ_i 所对应的那个特征向量，也有 $a_{nn}^{(k+1)}\to\lambda_i$, $a_{ln}^{(k+1)}=a_{nl}^{(k+1)}\to0$, $l=1,2,\cdots,n-1$.

上述推论都是假定使 $1/|\lambda_i-\sigma|$ 达到最大的特征值和使 $|\lambda_i-\sigma|$ 达到最大的特征值都只有一个，所对应的线性无关的特征向量也只有一个，才获得的. 从乘幂法和反迭代法得到的结果，收敛速度都是线性的，因此并没有显示 QL 方法的内在优点.

§2 用于对称三对角矩阵时的QL方法的性质

我们的兴趣是在将 QL 方法用在对称三对角矩阵上，因此我们要了解 QL 方法对于对称三对角矩阵有那些重要性质．下面记

$$T=\begin{pmatrix} \alpha_1 & \beta_1 & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1} \\ 0 & & & \beta_{n-1} & \alpha_n \end{pmatrix}$$

是对称三对角矩阵．

$$T-\sigma I=QL,$$

$$\hat{T}=LQ+\sigma I=\begin{pmatrix} \hat{\alpha}_1 & \hat{\beta}_1 & & & 0 \\ \hat{\beta}_1 & \hat{\alpha}_2 & \hat{\beta}_2 & & \\ & \hat{\beta}_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \hat{\beta}_{n-1} \\ 0 & & & \hat{\beta}_{n-1} & \hat{\alpha}_n \end{pmatrix}.$$

定理 3.5 若 T 不可约，则 L 的对角元 l_{22}, l_{33}, $\cdots$, l_{nn} 全不为 0.

证明 因为 T 不可约，故 β_1, β_2, $\cdots$, β_{n-1} 全不为 0. $T-\sigma I$ 的第 n 列

$$t_n=(0,\ 0,\ \cdots,\ 0,\ \beta_{n-1},\ \sigma_n-\sigma)^T\neq 0,$$

故
$$l_{nn}=\|t_n\|\neq 0,$$

T 的第 $n-1$ 列 $t_{n-1}=(0,\ 0,\ \cdots,\ 0,\ \beta_{n-2},\ \alpha_{n-1}-\sigma,\ \beta_{n-1})^T$，而

$$t_{n-1}-l_{n,n-1}q_n$$

的第 $n-2$ 个元素为 $\beta_{n-2}\neq 0$，故

$$l_{n-1,n-1}=\|t_{n-1}-l_{n,n-1}q_n\|\neq 0,$$

以此类推可得 $l_{n-2,n-2}\neq 0$，$l_{n-3,n-3}\neq 0$，…，$l_{33}\neq 0$，$l_{22}\neq 0$。证毕。

推论 T 不可约，$l_{11}=0$ 的充要条件是：σ 是 T 的特征值。

定理 3.5 和它的推论加强了定理 3.1 的结果。

定理 3.6 若 T 不可约，则

$$\hat{\beta}_i=\frac{l_{ii}}{l_{i+1,i+1}}\beta_i. \tag{7}$$

证明

$$q_n=\frac{1}{l_{nn}}\begin{pmatrix}0\\ \vdots\\ 0\\ \beta_{n-1}\\ \alpha_n-\sigma\end{pmatrix},$$

因此 q_n 的第 $n-1$ 个分量为 β_{n-1}/l_{nn}，第 1 个到第 $n-2$ 个分量全为 0。由

$$l_{kk}q_k=t_k-\sum_{j=k+1}^{n}l_{jk}q_j,\quad k=2,\ \cdots,\ n-1,$$

可知 q_k 的第 $k-1$ 个分量为 β_{k-1}/l_{kk}，第 1 个到第 $k-2$ 个分量全为 0。

由
$$\hat{T}=LQ+\sigma I,$$

$\hat{T}$ 的第 i 行第 $i+1$ 列为 $\hat{\beta}_i$ 是 L 的第 i 行与 Q 的第 $i+1$ 列 q_{i+1} 相乘，乘积为 $l_{ii}\beta_i/l_{i+1,i+1}$，故

$$\hat{\beta}_i=\frac{l_{ii}}{l_{i+1,i+1}}\beta_i,\quad i=1,\ 2,\ \cdots,\ n-1.$$ 证毕。

推论 1 T 不可约，且 σ 不是 T 的特征值，则 $\hat{T}$ 也是不可约.

推论 2 T 不可约，则 $\hat{\beta}_2, \hat{\beta}_3, \cdots, \hat{\beta}_{n-1}$ 不为 0，而 $\hat{\beta}_1=0$ 的充要条件是 σ 为 T 的特征值.

从(7)知 $\hat{\beta}_1$ 是 σ 的连续函数，因此当 σ 充分接近特征值时，$|\hat{\beta}_1|$ 就会很小. 而当 $|\hat{\beta}_1|$ 很小时，$\hat{\alpha}_1$ 就接近 T 的一个特征值. 后面这句话可以证明如下：

记 T 的特征值为 $\theta_1, \theta_2, \cdots, \theta_n$，由相似性 $\hat{T}$ 的特征值也是 $\theta_1, \theta_2, \cdots, \theta_n$. 记

$$\tilde{T}=\begin{pmatrix} \hat{\alpha}_1 & 0 & & & \\ 0 & \hat{\alpha}_2 & \hat{\beta}_2 & & \\ & \hat{\beta}_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \hat{\beta}_{n-1} \\ & & & \hat{\beta}_{n-1} & \hat{\alpha}_n \end{pmatrix},$$

它有一个特征值 $\hat{\alpha}_1$. 而

$$\|\hat{T}-\tilde{T}\|=\left\|\begin{pmatrix} 0 & \hat{\beta}_1 & & & \\ \hat{\beta}_1 & 0 & 0 & & \\ & 0 & \ddots & \ddots & \\ & & \ddots & \ddots & 0 \\ & & & 0 & 0 \end{pmatrix}\right\|=|\hat{\beta}_1|,$$

由对称矩阵的特征值摄动定理[文献 9, p. 462]，知可找到 $\hat{T}$ 的一个特征值 θ_i，使

$$|\hat{\alpha}_1-\theta_i|\leqslant\|\hat{T}-\tilde{T}\|=|\hat{\beta}_1|. \tag{8}$$

利用 Wielandt-Hoffman 定理[文献 9, p. 452]更有

$$(\hat{\alpha}_1-\theta_i)^2+(\tilde{\theta}_1-\theta_{i_1})^2+\cdots+(\tilde{\theta}_{n-1}-\theta_{i_{n-1}})^2\leqslant 2\beta_1, \tag{9}$$

这里 $\tilde{\theta}_1, \cdots, \tilde{\theta}_{n-1}$ 是 $\hat{T}_{2,n}$ 的特征值，$i_1, i_2, \cdots, i_{n-1}$ 是 $1, 2, \cdots, i-1, i+1, \cdots, n$ 的一种排列.

(8)、(9) 两式告诉我们，对于对称三对角矩阵 T，使用带位移 $\{\sigma_k\}$ 的 QL 方法，求特征值，得到 $\{T^{(k)}\}$，如果 $T^{(k)}$ 的 β_1 充分小时，它的 α_1 即为一个近似特征值，接着只要对 $T^{(k)}_{2,n}$ 继续进行 QL 方法，求其余的特征值. 这样每求一个特征值，还可把矩阵降低一阶，求到的特征值还可以用 (8) 式和 (9) 式来估计误差，这就显示出 QL 方法的优点. 这里也可看出 (8) 和 (9) 那样的估计式的意义. 如果能改进 (8) 式、(9) 式的结果，是很有实用价值的.

W. Kahan 在 1966 年曾考虑过如下的问题：对称三对角矩阵 T 中某一个 β_i 用 0 代替后的矩阵为 $\tilde{T}$，它的特征值为 $\tilde{\theta}_1, \tilde{\theta}_2, \cdots, \tilde{\theta}_n$，问 $\sum\limits_{i=1}^{n}(\theta_i-\tilde{\theta}_i)^2$ 能否得到比 (9) 更好一点的估计式. 他给出了一个结果

$$\sum_{i=1}^{n}(\theta_i-\tilde{\theta}_i)^2\leqslant\frac{\beta_i^2}{\delta_i^2+\rho_i^2}\left[2\rho_i^2+\frac{\delta_i^2\beta_i^2}{\delta_i^2+\rho_i^2}\right],$$

其中

$$\delta_i=(\alpha_{i+1}-\alpha_i)/2,\quad \rho_i^2=(1-1/\sqrt{2})\,(\beta_{i-1}^2+\beta_{i+1}^2),$$

见文献 [10, p. 134].

对于 (8) 式至少可有

定理 3.7 对称三对角矩阵 T 的特征值为 $\theta_1, \theta_2, \cdots, \theta_n$，

$$\theta_1\leqslant\theta_2\leqslant\theta_3\leqslant\cdots\leqslant\theta_n,$$

T 的某个 β_i 用 0 代替后的新矩阵 $\tilde{T}$，它的特征值为 $\tilde{\theta}_1, \tilde{\theta}_2, \cdots, \tilde{\theta}_n$，$\tilde{\theta}_1\leqslant\tilde{\theta}_2\leqslant\cdots\leqslant\tilde{\theta}_n$，假定 $\beta_{i-1}^2+\beta_{i+1}^2\neq 0$，则

$$|\theta_i-\tilde{\theta}_i|<|\beta_i|. \tag{10}$$

证明 不失一般性假定 $\beta_i>0$，令

$$T=\tilde{T}+E,$$

$$E=\begin{pmatrix} 0 & 0 & & & & \\ & \ddots & \ddots & & & \\ 0 & 0 & \beta_i & & & \\ & \ddots & \ddots & & & \\ & \beta_i & \ddots & 0 & & \\ & & & \ddots & \ddots & \\ & & 0 & \ddots & 0 & \\ & & & \ddots & \ddots & \\ & & & 0 & 0 \end{pmatrix} \begin{matrix} \text{第 } i \text{ 行} \\ \text{第 } i+1 \text{ 行} \end{matrix},$$

易知 E 的特征值依小到大次序为 $-\beta_i$, 0, 0, …, 0, β_i, 记对应的特征向量为 $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n$, 易知

$$\boldsymbol{b}_1=\frac{1}{\sqrt{2}}(\boldsymbol{e}_i-\boldsymbol{e}_{i+1}),$$

$$\boldsymbol{b}_n=\frac{1}{\sqrt{2}}(\boldsymbol{e}_i+\boldsymbol{e}_{i+1}),$$

记 $\widetilde{T}$ 的特征值 $\tilde{\theta}_j$ 对应的特征向量为 $\boldsymbol{x}_j$. 利用极大极小原理(见[9]或本书第 4 章定理 4.1)有

$$\theta_j=\max_{V_{n-j+1}}\min_{x\in V_{n-j+1}}\rho(\boldsymbol{x}, T),$$

这里 $\rho(x, B)=(Bx, x)/(x, x)$, V_{n-j+1} 是一个 $n-j+1$ 维的线性子空间. 于是

$$\theta_j\geqslant\min_{\boldsymbol{x}\in\{\boldsymbol{x}_j, \boldsymbol{x}_{j+1}, \cdots, \boldsymbol{x}_n\}}\rho(\boldsymbol{x}, T),$$

右边的极小是能达到的, 记在 $\tilde{\boldsymbol{x}}$ 处达到, 可设 $\|\tilde{\boldsymbol{x}}\|=1$,

$$\theta_j\geqslant\rho(\tilde{\boldsymbol{x}}, T)=\rho(\tilde{\boldsymbol{x}}, \widetilde{T})+\rho(\tilde{\boldsymbol{x}}, E).$$

对于 $\tilde{\boldsymbol{x}}$ 可分如下两种情况:

1. 若 $\tilde{\boldsymbol{x}}$ 是 $\widetilde{T}$ 的对应特征值 $\tilde{\theta}_j$ 的特征向量, 此时

$$\rho(\tilde{\boldsymbol{x}}, \widetilde{T})=\tilde{\theta}_j,$$

下面证明此时有 $\rho(\tilde{\boldsymbol{x}}, E)>-\beta_i$, 实际上因为 $\tilde{\boldsymbol{x}}\neq\pm b_1$, 否则 $\boldsymbol{b}_1$ 是 $\widetilde{T}$ 的对应 $\tilde{\theta}_j$ 的特征向量, 但

$(\tilde{T}-\tilde{\theta}_j)\boldsymbol{b}_1$

$$=\frac{1}{\sqrt{2}}(0,\ \cdots,\ 0,\ \underset{\text{第 } i-1 \text{ 列}}{\beta_{i-1}},\ *,\ *,\ \underset{\text{第 } i+2 \text{ 列}}{-\beta_{i+1}},\ 0,\ 0,\ \cdots,\ 0)^T\neq 0,$$

因此$(\tilde{\boldsymbol{x}},\ \boldsymbol{b}_1)^2=h_1^2<1$.

若 $\tilde{\boldsymbol{x}}=\sum\limits_{l=1}^{n}h_l\boldsymbol{b}_l$,

$$\rho(\tilde{\boldsymbol{x}},\ E)=-\beta_i h_1^2+h_n^2\beta_i>-\beta_i.$$

2. $\tilde{\boldsymbol{x}}$ 不是 $\tilde{T}$ 的对应 $\tilde{\theta}_j$ 的特征向量,此时

$$\rho(\tilde{x},\ \tilde{T})>\tilde{\theta}_j,\quad \rho(\tilde{x},\ E)\geqslant-\beta_i,$$

因此不管哪种情况都有

$$\theta_j>\tilde{\theta}_j-\beta_i.$$

又利用

$$\theta_j=\min_{V_j}\max_{x\in V_j}\rho(\boldsymbol{x},\ T)\leqslant\max_{x\in\{x_1,x_2,\cdots,x_j\}}\rho(\boldsymbol{x},\ T),$$

同样可证得：$\theta_j<\tilde{\theta}_j+\beta_i$, 故

$$|\theta_j-\tilde{\theta}_j|<\beta_i.$$

从估计式(8)、(9)、(10)可知 $\{T^{(k)}\}$ 中 $T^{(k)}$ 的 β_1 趋于 0 的速度快,那么 α_1 趋于 T 的某个特征值的速度也快.

下面我们再给出一个估计 $\hat{\beta}_1$ 的关系式(见文献[10]).

定理 3.8 设 $T-\sigma I=QL,\hat{T}=LQ+\sigma I,Q=[\boldsymbol{q}_1,\boldsymbol{q}_2,\cdots,\boldsymbol{q}_n]$是正交阵,$L$ 是下三角阵,记 $\boldsymbol{q}_1$ 与 $\boldsymbol{e}_1$ 的夹角为 θ,则

$$(T-\sigma I)\boldsymbol{q}_1=l_{11}\boldsymbol{e}_1,\tag{11}$$

$$|\hat{\beta}_1|=l_{11}|\sin\theta|.\tag{12}$$

证明 $T-\sigma I=QL$,

$T-\sigma I=L^TQ^T$,

$(T-\sigma I)Q=L^T$,

故 $$(T-\sigma I)\boldsymbol{q}_1=l_{11}\boldsymbol{e}_1,$$

(11)式获证. 又

$$Q\hat{T}=TQ,$$

故
$$\hat{\alpha}_1\boldsymbol{q}_1+\hat{\beta}_1\boldsymbol{q}_2=T\boldsymbol{q}_1,$$

$$\begin{aligned}\hat{\beta}_1\boldsymbol{q}_2&=T\boldsymbol{q}_1-\hat{\alpha}_1\boldsymbol{q}_1=T\boldsymbol{q}_1-\boldsymbol{q}_1^TT\boldsymbol{q}_1\boldsymbol{q}_1=(I-\boldsymbol{q}_1\boldsymbol{q}_1^T)T\boldsymbol{q}_1\\&=(I-\boldsymbol{q}_1\boldsymbol{q}_1^T)(l_{11}\boldsymbol{e}_1+\sigma\boldsymbol{q}_1)\\&=l_{11}(I-\boldsymbol{q}_1\boldsymbol{q}_1^T)\boldsymbol{e}_1=l_{11}(\boldsymbol{e}_1-\boldsymbol{q}_1\cos\theta),\end{aligned}$$

于是
$$\hat{\beta}_1^2=l_{11}^2(1-2\cos^2\theta+\cos^2\theta)=l_{11}^2\sin^2\theta,$$

故(12)成立. 证毕.

推论 $|\hat{\beta}_1|\leqslant l_{11}$.

由定理 3.8 可知, 为了估计 β_1 的收敛速度, 估计 l_{11} 是很重要的一环. 我们利用第 2 章定理 2.6 的结果给出如下的 l_{11} 和 $\hat{\beta}_1$ 的表示式, 参见[31].

定理 3.9 设 T 不可约, 则

$$l_{11}^2=\det(T-\sigma I)^2/\omega_1, \tag{13}$$

其中

$$\begin{aligned}\omega_1=&\det(T_{2n}-\sigma I)^2+(\det(T_{3,n}-\sigma I)\beta_1)^2\\&+(\det(T_{4,n}-\sigma I)\beta_1\beta_2)^2+\cdots+(\beta_1\beta_2\cdots\beta_{n-1})^2.\end{aligned} \tag{14}$$

证明 先设 σ 不是 T 的特征值, 故 $\det(T-\sigma I)\neq 0$, 由(11)$T-\sigma I$ 是一个对称三对角阵, 因此 $\boldsymbol{q}_1$ 的分量 q_{l1} ($l=1, 2, \cdots, n$), 可以利用定理 2.6 有表示式

$$q_{l1}=(-1)^{l-1}l_{11}\frac{\beta_1\beta_2\cdots\beta_{l-1}\det(T_{l+1,n}-\sigma I)}{\det(T-\sigma I)}, \tag{15}$$

但 $\sum\limits_{l=1}^{n}q_{l1}^2=1$, 故

$$\begin{aligned}&l_{11}^2(\det(T_{2,n}-\sigma I)^2+(\det(T_{3,n}-\sigma)\beta_1)^2+\cdots\\&+(\beta_1\beta_2\cdots\beta_{n-1})^2)/\det(T-\sigma I)^2=1,\end{aligned}$$

因为 T 不可约, 故 $\omega_1>0$, 于是即得(13).

如果 σ 是 T 的特征值，则知 $l_{11}=0$, $\det(T-\sigma I)=0$, (13) 自然成立。证毕.

定理 3.10 设 T 不可约，令 ω_1 如定理 3.9 中所定义，

$$\omega_2=\det(T_{3,n}-\sigma I)^2+(\det(T_{4,n}-\sigma I)\beta_2)^2+\cdots+(\beta_2\beta_3\cdots\beta_{n-1})^2,$$

则

$$\hat{\beta}_1^2=\beta_1^2\det(T-\sigma I)^2\omega_2/\omega_1^2. \tag{16}$$

证明 由定理 3.8 知

$$\hat{\beta}_1^2=l_{11}^2\sin^2\theta,$$

但 $q_{11}=\cos\theta$, $\sin^2\theta=1-q_{11}^2=\sum_{l=2}^{n}q_{l1}^2$, 故由(15)

$$\sin^2\theta=l_{11}^2\sum_{l=2}^{n}(\beta_1\beta_2\cdots\beta_{l-1}\det(T_{l+1,n}-\sigma I))^2/\det(T-\sigma I)^2$$

$$=l_{11}^2\beta_1^2\omega_2/\det(T-\sigma I)^2=\beta_1^2\omega_2/\omega_1,$$

故

$$\hat{\beta}_1^2=l_{11}^2\sin^2\theta=\frac{\det(T-\sigma I)^2}{\omega_1}\times\frac{\beta_1^2\omega_2}{\omega_1}$$

$$=\beta_1^2\det(T-\sigma I)^2\omega_2/\omega_1^2.$$ 证毕.

(16) 式给 QL 方法的收敛速度估计提供了非常有效的公式. 由此可知 $\hat{\beta}_1$ 收敛于 0 的速度，依赖于 $\det(T-\sigma I)$ 收敛于 0 的速度. 对于不同的位移 σ_k, 就有不同的 $\det(T-\sigma_k I)$ 收敛状况，本章的最后两节要介绍两种最常用的位移$\{\sigma_k\}$的取法，并分析带那种位移时，QL 方法的收敛状况.

在研究收敛性前，我们先介绍一下：用旋转阵逐次相乘来实现 QL 迭代的计算方法；它可以有递推的计算公式，并且可以节省计算量. 对于对称三对角矩阵 T, 和位移 σ, $T-\sigma I=QL$, $\hat{T}=LQ+\sigma I$, 我们要给出从 T 求 $\hat{T}$ 的计算公式，也即从 $\alpha_1, \alpha_2, \cdots, \alpha_n, \beta_1, \beta_2, \cdots, \beta_{n-1}$ 求 $\hat{\alpha}_1, \hat{\alpha}_2, \cdots, \hat{\alpha}_n, \hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_{n-1}$ 的计算公式. 记 $\bar{a}_k=\alpha_k-\sigma$, $k=1, 2, \cdots, n$, 即

$$T-\sigma I=\begin{pmatrix} \bar{\alpha}_1 & \beta_1 & & & 0 \\ \beta_1 & \bar{\alpha}_2 & \beta_2 & & \\ & \beta_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1} \\ 0 & & & \beta_{n-1} & \alpha_n \end{pmatrix},$$

将它左乘旋转阵 $R_{n-1,n}$，将 $(n-1,\ n)$ 位置上的元素化成零.

$$\begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ & & c_n & -s_n \\ 0 & & s_n & c_n \end{pmatrix}\begin{pmatrix} \bar{\alpha}_1 & \beta_1 & & & 0 \\ \beta_1 & \bar{\alpha}_2 & \ddots & & \\ & \ddots & \ddots & \beta_{n-2} & \\ & & \beta_{n-2} & \bar{\alpha}_{n-1} & \beta_{n-1} \\ 0 & & & \beta_{n-1} & \bar{\alpha}_n \end{pmatrix}$$

$$=\begin{pmatrix} \bar{\alpha}_1 & \beta_1 & & & & \\ \beta_1 & \bar{\alpha}_2 & \beta_2 & & & \\ & \beta_2 & \ddots & \ddots & & \\ & & \ddots & \ddots & \beta_{n-3} & \\ & & & \beta_{n-3} & \bar{\alpha}_{n-2} & \beta_{n-2} \\ & & & & c_n\beta_{n-2} & c_n\bar{\alpha}_{n-1}-s_n\beta_{n-1} & c_n\beta_{n-1}-s_n\bar{\alpha}_n \\ & & & & s_n\beta_{n-2} & s_n\bar{\alpha}_{n-1}+\beta_{n-1}c_n & s_n\beta_{n-1}+c_n\bar{\alpha}_n \end{pmatrix},$$

为了使 $(n-1,\ n)$ 位置上的元素为 0，只要 $c_n\beta_{n-1}-s_n\bar{\alpha}_n=0$，即得

$$c_n=\bar{\alpha}_n/r_n,\ s_n=\beta_{n-1}/r_n,\ r_n=\sqrt{\bar{\alpha}_n^2+\beta_{n-1}^2},$$

记 $t_n=s_n\beta_{n-2},\ q_n=s_n\bar{\alpha}_{n-1}+\beta_{n-1}c_n,\ p_{n-1}=c_n\bar{\alpha}_{n-1}-s_n\beta_{n-1},$

于是上面等式右边的矩阵是

$$\begin{pmatrix} \bar{\alpha}_1 & \beta_1 & & & & \\ \beta_1 & \bar{\alpha}_2 & \ddots & & 0 & \\ & \ddots & \ddots & \beta_{n-3} & & \\ & & \beta_{n-3} & \bar{\alpha}_{n-2} & \beta_{n-2} & \\ & & & c_n\beta_{n-2} & p_{n-1} & 0 \\ & 0 & & t_n & q_n & r_n \end{pmatrix},$$

对 $R_{n-1,n}(T-\sigma I)$ 再左乘旋转阵 $R_{n-2,n-1}$，将 $(n-2, n-1)$ 位置上的元素化成零.

$$R_{n-2,n-1}R_{n-1,n}(T-\sigma I)$$

$$=\begin{pmatrix} \bar{\alpha}_1 & \beta_1 & & & & \\ \beta_1 & \ddots & & & & \\ & & \bar{\alpha}_{n-3} & \beta_{n-3} & & \\ & & c_{n-1}\beta_{n-3} & p_{n-2} & 0 & \\ & & t_{n-1} & q_{n-1} & r_{n-1} & \\ & & & t_n & q_n & r_n \end{pmatrix},$$

其中

$$\begin{cases} r_{n-1}=\sqrt{\beta_{n-2}^2+p_{n-1}^2}, \\ s_{n-1}=\beta_{n-2}/r_{n-1}, \ c_{n-1}=p_{n-1}/r_{n-1}, \\ p_{n-2}=c_{n-1}\bar{\alpha}_{n-2}-s_{n-1}c_n\beta_{n-2}, \\ q_{n-1}=s_{n-1}\bar{\alpha}_{n-2}+c_{n-1}c_n\beta_{n-2}, \\ t_{n-1}=s_{n-1}\beta_{n-3}. \end{cases} \tag{17}$$

将(17)式与前面 $R_{n-1,n}$ 旋转时的量进行对比，就会发现，如果令 $\bar{\alpha}_n=p_n$, $c_{n+1}=1$，就有一般的递推公式：

$$\begin{cases} r_k=(\beta_{k-1}^2+p_k^2)^{\frac{1}{2}}, \\ s_k=\beta_{k-1}/r_k,\ c_k=p_k/r_k, \\ p_{k-1}=c_k\bar{\alpha}_{k-1}-s_kc_{k+1}\beta_{k-1}, \\ q_k=s_k\bar{\alpha}_{k-1}+c_kc_{k+1}\beta_{k-1}, \\ t_k=s_k\beta_{k-2},\quad k=n,\ n-1,\ \cdots,\ 2. \end{cases} \tag{18}$$

得到

$$R_{1,2}R_{2,3}\cdots R_{n-2,n-1}R_{n-1,n}(T-\sigma I)$$

$$=\begin{pmatrix} p_1 & & & & & 0 \\ q_2 & r_2 & & & & \\ t_3 & q_3 & r_3 & & & \\ & \ddots & \ddots & \ddots & & \\ & & & & & \\ 0 & & & t_n & q_n & r_n \end{pmatrix}=L.$$

同时知道

$$Q=R_{n-1,n}^*R_{n-2,n-1}^*\cdots R_{2,3}^*R_{1,2}^*.$$

现在来考察 $LQ=LR_{n-1,n}^*R_{n-2,n-1}^*\cdots R_{2,3}^*R_{1,2}^*$, $LR_{n-1,n}^*$ 只改变 L 的最后两列；而 $LR_{n-1,n}^*$ 再右乘 $R_{n-2,n-1}^*\cdots R_{2,3}^*R_{1,2}^*$，不改变最后一列，因此从 $LR_{n-1,n}^*$ 的最后一列可以得到 $\hat{\beta}_{n-1}$ 和 $\hat{\alpha}_n$。通过简单的相乘得到

$$\begin{cases} \hat{\beta}_{n-1}=r_{n-1}s_n, \\ \hat{\alpha}_n=q_ns_n+r_nc_n+\sigma, \end{cases} \tag{19}$$

$$LR^*_{n-1,n}=\begin{pmatrix} p_1 & & & & & 0 \\ q_2 & r_2 & & & & \\ t_3 & \ddots & \ddots & & & \\ & \ddots & \ddots & r_{n-2} & & \\ & & \ddots & q_{n-1} & r_{n-1}c_n & \hat{\beta}_{n-1} \\ 0 & & & t_n & x & q_ns_n+r_nc_n \end{pmatrix},$$

对 $LR^*_{n-1,n}$ 再右乘 $R^*_{n-2,n-1}$, 只改变第 $n-2$ 列和 $n-1$ 列, 得到

$$\begin{cases} \hat{\beta}_{n-2}=r_{n-2}s_{n-1}, \\ \hat{\alpha}_{n-1}=q_{n-1}s_{n-1}+r_{n-1}c_nc_{n-1}+\sigma, \end{cases} \tag{20}$$

在 $LR^*_{n-1,n}R^*_{n-2,n-1}$ 中 $(n-2,n-2)$ 位置上的元素变成 $c_{n-1}r_{n-2}$, 比较(19)与(20), 如果考虑到 $c_{n+1}=1$, 再令 $p_1=r_1$, 则有一般的递推公式

$$\left.\begin{array}{l} \hat{\beta}_k=r_ks_{k+1}, \\ \hat{\alpha}_{k+1}=q_{k+1}s_{k+1}+r_{k+1}c_{k+2}c_{k+1}+\sigma, \\ k=n-1,\ n-2,\ \cdots,\ 1, \\ \hat{\alpha}_1=r_1c_2+\sigma, \end{array}\right\}$$

考虑到 $r_{k+1}c_{k+1}=p_{k+1}$, 上式可以写成

$$\begin{cases} \hat{\beta}_k=r_ks_{k+1}, \\ \hat{\alpha}_{k+1}=q_{k+1}s_{k+1}+p_{k+1}c_{k+2}+\sigma, \\ k=n-1,\ n-2,\ \cdots,\ 1, \\ \hat{\alpha}_1=r_1c_2+\sigma. \end{cases} \tag{21}$$

(18)式与(21)式合并在一起, 就是从 T 计算 $\hat{T}$ 的计算公式, 在(18)式中的 t_k 这个量, 在(21)式中没有用到, 在(18)式中的其他量中也没有用到它, 因此可以不必计算. 这样总共需:

加减法运算　　$5n-4$ 次，

乘法运算　　$10n-9$ 次，

除法运算　　$2n-2$ 次，

开方运算　　$n-1$ 次.

一般计算机对开方运算总是多费一点时间，因此很多数值分析专家要想法把开方运算去掉，对此提出不少算法，其中 C. H. Reinsch 在 1971 年发表的文章 [11] 中的算法是非常成功的. 他是对 QR 方法来说的，我们将它修改成对 QL 方法使用的，介绍如下：在算式(18)和(21)中引进二个新的量

$$h_k=p_kc_{k+1},\quad g_k=p_k/c_{k+1},$$

有 $h_n=g_n=p_n=\bar{\alpha}_n$，再注意到成立等式

$$c_k\beta_{k-1}=s_kp_k,$$

于是从(18)式

$$p_{k-1}=c_k\bar{\alpha}_{k-1}-s_kc_{k+1}\beta_{k-1},$$

可得

$$\begin{aligned}g_{k-1}&=\bar{\alpha}_{k-1}-s_kc_{k+1}\beta_{k-1}/c_k=\bar{\alpha}_{k-1}-s_kc_{k+1}\beta_{k-1}^2/c_k\beta_{k-1}\\&=\bar{\alpha}_{k-1}-s_kc_{k+1}\beta_{k-1}^2/s_kp_k=\bar{\alpha}_{k-1}-\beta_{k-1}^2/g_k,\end{aligned}\tag{22}$$

这是一个从 g_k 算 g_{k-1} 的递推公式，式中没有开方运算. 又

$$h_{k-1}=g_{k-1}p_k^2/r_k^2,\tag{23}$$

也即知道 p_k、r_k 后可以算 h_{k-1}. 而

$$r_k^2=p_k^2+\beta_{k-1}^2,\tag{24}$$

$$p_{k-1}^2=g_{k-1}h_{k-1},\tag{25}$$

这些都是没有开方运算的算式，它们告诉我们从 T 可以不经开方运算，得到 r_k, p_k, h_k, g_k, 因此下面只要将 $\hat{\beta}_k$ 和 $\hat{\alpha}_k$ 用这些量和 T 的元素表示出来，表示式中要求不包含开方运算. 从(21)

$$
\begin{aligned}
\hat{\alpha}_{k+1} &= q_{k+1}s_{k+1} + p_{k+1}c_{k+2} + \sigma \\
&= h_{k+1} + \sigma + (s_{k+1}\bar{\alpha}_k + c_{k+1}c_{k+2}\beta_k)s_{k+1} \\
&= h_{k+1} + \sigma + s_{k+1}^2(\bar{\alpha}_k + h_{k+1}),
\end{aligned}
$$

$$
\hat{\alpha}_1 = r_1c_2 + \sigma = h_1 + \sigma,
$$

$$
\hat{\beta}_k^2 = r_k^2 s_{k+1}^2,
$$

综合这些式子得到无开方运算的 QL 迭代的算式，

初始 $p_n = g_n = h_n = \bar{\alpha}_n,\quad r_n^2 = p_n^2 + \beta_{n-1}^2,$

$$
\begin{cases}
s_k^2 = \beta_{k-1}^2 / r_k^2, \\
\hat{\alpha}_k = h_k + \sigma + s_k^2(\bar{\alpha}_{k-1} + h_k), \\
g_{k-1} = \bar{\alpha}_{k-1} - \beta_{k-1}^2 / g_k, \\
h_{k-1} = g_{k-1}p_k^2 / r_k^2,\ p_{k-1}^2 = g_{k-1}h_{k-1}, \\
r_{k-1}^2 = p_{k-1}^2 + \beta_{k-2}^2,\ \beta_0^2 = 0, \\
\hat{\beta}_{k-1}^2 = r_{k-1}^2 s_k^2, \\
k = n,\ n-1,\ \cdots,\ 2, \\
\hat{\alpha}_1 = h_1 + \sigma.
\end{cases}
\tag{26}
$$

这是从 α_1、α_2、$\cdots$、α_n、β_1^2、β_2^2、$\cdots$、β_{n-1}^2 算出 $\hat{\alpha}_1$、$\hat{\alpha}_2$、$\cdots$、$\hat{\alpha}_n$、$\hat{\beta}_1^2$、$\hat{\beta}_2^2$、$\cdots$、$\hat{\beta}_{n-1}^2$ 的公式. 尽管算出的不是 $\hat{\beta}_k$ 而是 $\hat{\beta}_k^2$，但是对 QL 方法来说，这是没有妨害的.

不计从 $\beta_1,\ \beta_2,\ \cdots,\ \beta_{n-1}$ 得到 $\beta_1^2,\ \beta_2^2, \cdots, \beta_{n-1}^2$ 的计算量，总的计算量是

加减法运算：$5n-3$ 次，

除法运算：　$3n-3$ 次，

乘法运算：　$4n-3$ 次.

在计算过程中遇到 $g_i \simeq 0$ 时，可以用小的正数 δ 来代替，计算可以继续进行，从 g_i 表示式中可知这样的代替相当于对 α_i 有一个摄动，变成 $\alpha_i + \delta$；只要 δ 充分小，仍可使求得的特征值满足预定的精度要求.

§3 带 Rayleigh 商位移的 QL 方法

对于带位移$\{\sigma_k\}$的 QL 方法:

$$T^{(k)}-\sigma_k I=Q_k L_k,$$
$$T^{(k+1)}=L_k Q_k+\sigma_k I,$$

如果

$$T^{(1)}=T,$$

$$T^{(k)}=\begin{pmatrix} \alpha_1^{(k)} & \beta_1^{(k)} & & & 0 \\ \beta_1^{(k)} & \alpha_2^{(k)} & \beta_2^{(k)} & & \\ & \beta_2^{(k)} & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1}^{(k)} \\ 0 & & & \beta_{n-1}^{(k)} & \alpha_n^{(k)} \end{pmatrix},$$

而取 $\alpha_1^{(k)}$ 作为 σ_k, 那么这时的 QL 方法, 称为带 Rayleigh 商位移的 QL 方法.

为了弄清为何叫做 Rayleigh 商位移, 先要了解什么是 Rayleigh 商. 对于任意对称矩阵 A, 取向量 $\boldsymbol{x}\neq 0$, 那么

$$\rho(\boldsymbol{x})=(A\boldsymbol{x},\ \boldsymbol{x})/(\boldsymbol{x},\ \boldsymbol{x})$$

称为 A 在向量 $\boldsymbol{x}$ 时的 Rayleigh 商, 它当然跟 A 有关, 因此有时为了强调一下对应的矩阵, 记为 $\rho(\boldsymbol{x},\ A)$. 由对称矩阵的极大极小原理知道 $\rho(\boldsymbol{x})$ 的极值就是 A 的特征值, 因此可以通过 Rayleigh 商来求得 A 的特征值.

所谓 Rayleigh 商迭代法, 就是一种求对称矩阵特征值和对应特征向量的方法. 算法如下:

1. 取一个单位向量 $\boldsymbol{x}_1$, $1\to k$;
2. 计算 $\rho_k=\rho(\boldsymbol{x}_k)$;
3. 如果 $\det(A-\rho_k I)\neq 0$, 则从 $(A-\rho_k I)\boldsymbol{y}_{k+1}=\boldsymbol{x}_k$, 求

$\boldsymbol{y}_{k+1}$，再将 $\boldsymbol{y}_{k+1}$ 标准化成为 $\boldsymbol{x}_{k+1}$，$\boldsymbol{x}_{k+1}=\boldsymbol{y}_{k+1}/\|\boldsymbol{y}_{k+1}\|$，如果 $\|\boldsymbol{y}_{k+1}\|$ 很大，ρ_k 作为 A 的特征值近似值，$\boldsymbol{x}_{k+1}$ 作为对应特征向量的近似，计算完成。

如果 $\|\boldsymbol{y}_{k+1}\|$ 不很大，则 $k+1\to k$ 转 2。

如果 $\det(A-\rho_k I)=0$，从

$$(A-\rho_k I)\boldsymbol{x}_{k+1}=0$$

算得单位向量 $\boldsymbol{x}_{k+1}$，ρ_k、$\boldsymbol{x}_{k+1}$ 是 A 的特征值和对应的特征向量，计算完成。

对于 Rayleigh 商迭代法，是否从任意 $\boldsymbol{x}_1$ 出发，ρ_k 都是收敛到 A 的一个特征值？这就是 Rayleigh 商迭代法的收敛性问题．我们先不来回答这个问题．还是再来回答那样的位移，为何叫带 Rayleigh 商位移．有如下定理[10]．

定理 3.11 对 $T^{(1)}=T$ 按带 Rayleigh 位移的 QL 方法获得矩阵序列 $\{T^{(k)}\}$，和从 $\boldsymbol{x}_1=\boldsymbol{e}_1$ 按 Rayleigh 迭代法获得的序列 $\{\rho_k\}$，成立如下关系

$$\rho_k=\alpha_1^{(k)}, \tag{27}$$

$$\boldsymbol{x}_{k+1}=P_k\boldsymbol{e}_1. \tag{28}$$

证明 因为

$$\rho_1=(T\boldsymbol{x}_1,\ \boldsymbol{x}_1)/(\boldsymbol{x}_1,\ \boldsymbol{x}_1)=(T\boldsymbol{e}_1,\ \boldsymbol{e}_1)/(\boldsymbol{e}_1,\ \boldsymbol{e}_1)=\alpha_1^{(1)},$$

因此(27)式对 $k=1$ 成立，又

$$(T-\rho_1 I)\boldsymbol{y}_2=\boldsymbol{x}_1=\boldsymbol{e}_1,$$

由定理 3.8 知 $\boldsymbol{y}_2$ 与 $\boldsymbol{q}_1^{(1)}$ 至多差一个正常数因子，即 $\boldsymbol{x}_2=\boldsymbol{q}_1^{(1)}=Q_1\boldsymbol{e}_1=P_1\boldsymbol{e}_1$，因此(28)式也对 $k=1$ 成立。

现在用数学归纳法，假定(27)、(28)对 $k=1,\ 2,\ \cdots,\ j$ 都成立，考虑 $k=j+1$ 的情况，此时

$$\begin{aligned}\rho_{j+1}&=(T\boldsymbol{x}_{j+1},\ \boldsymbol{x}_{j+1})/(\boldsymbol{x}_{j+1},\ \boldsymbol{x}_{j+1})\\&=(TP_j\boldsymbol{e}_1,\ P_j\boldsymbol{e}_1)=(P_j^*TP_j\boldsymbol{e}_1,\ \boldsymbol{e}_1),\end{aligned}$$

但由 QL 方法知

$$P_j^* T P_j = T^{(j+1)}$$

成立，于是 $\rho_{j+1} = (T^{(j+1)}\boldsymbol{e}_1,\ \boldsymbol{e}_1) = \alpha_1^{(j+1)}$,

又

$$(T-\rho_{j+1})\boldsymbol{y}_{j+2} = \boldsymbol{x}_{j+1},$$

$$(T-\alpha_1^{(j+1)})\boldsymbol{y}_{j+2} = P_j\boldsymbol{e}_1,$$

$$(P_j^* T P_j - \alpha_1^{(j+1)}) P_j^* \boldsymbol{y}_{j+2} = \boldsymbol{e}_1,$$

$$(T^{(j+1)} - \alpha_1^{(j+1)}) P_j^* \boldsymbol{y}_{j+2} = \boldsymbol{e}_1,$$

再由定理 3.8 知 $P_j^*\boldsymbol{y}_{j+1}$ 与 $\boldsymbol{q}_1^{(j+1)}$ 至多差一个正常数因子，即

$$P_j^*\boldsymbol{y}_{j+2} = \tau Q_{j+1}\boldsymbol{e}_1,$$

或 $\boldsymbol{y}_{j+2} = \tau P_j Q_{j+1}\boldsymbol{e}_1 = \tau P_{j+1}\boldsymbol{e}_1$,

从而 $\boldsymbol{x}_{j+2} = P_{j+1}\boldsymbol{e}_1$. 证毕.

根据这个定理可知，带 Rayleigh 商位移的 QL 方法中的 β_1 是否收敛于零跟从 $\boldsymbol{e}_1$ 出发的 Rayleigh 商迭代是否收敛，是等价的. 实际上，若 $\rho_k \to \theta_i$, θ_i 是 T 的一个特征值，$\boldsymbol{x}_k \to \boldsymbol{y}_i$, $\boldsymbol{y}_i$ 是对应 θ_i 的单位特征向量，于是按定理 3.11，P_j 的第 1 列趋于 $\boldsymbol{y}_i$，由

$$T^{(j+1)} = P_j^* T P_j = P_j^*[\theta_i \boldsymbol{y}_i + \varepsilon,\ T\boldsymbol{g}_2,\ \cdots,\ T\boldsymbol{g}_n]$$

知 $\beta_1^{(j+1)} \to 0$，这里 ε 是趋于 0 的向量，$\boldsymbol{g}_2$、$\cdots$、$\boldsymbol{g}_n$ 是正交矩阵 P_j 的第 2 列、…、第 n 列. 反过来也成立，如果 $\beta_1^{(k)} \to 0$，则 $\rho_k \to \theta_i$, $\boldsymbol{x}_{k+1} \to \boldsymbol{y}_i$.

自然人们会提出这样的问题：对任意对称三对角矩阵 T，使用带 Rayleigh 商位移的 QL 方法，$\beta_1^{(k)}$ 是否一定能收敛于零？下面的例子说明对有些矩阵 T，$\beta_1^{(k)}$ 是不收敛的.

例 $T = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}$ 使用带 Rayleigh 商位移的 QL 方法.

$$T^{(1)}=T,\ \sigma_1=\alpha_1^{(1)}=\alpha_1=0,\ T^{(1)}-\sigma_1 I=Q_1L_1,$$

$$Q_1=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad L_1=\begin{pmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{pmatrix},$$

$$T^{(2)}=L_1Q_1+\sigma_1 I=\begin{pmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}=\begin{pmatrix}0 & \frac{1}{2}\\ \frac{1}{2} & 0\end{pmatrix},$$

由此可知 $T^{(k)}=T^{(1)},\ k=2,\ 3,\ \cdots,\ \sigma_k=\alpha_1^{(k)}=0,$

$$\beta_1^{(k)}=\frac{1}{2}$$

不收敛于零。

这个例子当然是很特殊的情况.在这个情况下,Rayleigh 商迭代法,取 $\boldsymbol{x}_1=\boldsymbol{e}_1$, ρ_k 也是不能收敛到特征值的. 我们来考察一下,此时 ρ_k 和 $\boldsymbol{x}_k$ 的状况.

$$\boldsymbol{x}_1=\boldsymbol{e}_1,\ \rho_1=\rho(\boldsymbol{x}_1)=(T\boldsymbol{e}_1,\ \boldsymbol{e}_1)=0,\ (T-\rho_1 I)\boldsymbol{y}_2=\boldsymbol{x}_1,$$

即
$$\begin{pmatrix}0 & \frac{1}{2}\\ \frac{1}{2} & 0\end{pmatrix}\begin{pmatrix}y_{12}\\ y_{22}\end{pmatrix}=\begin{pmatrix}1\\ 0\end{pmatrix},$$

得到
$$\boldsymbol{y}_2=\begin{pmatrix}0\\ 2\end{pmatrix},\quad \boldsymbol{x}_2=\begin{pmatrix}0\\ 1\end{pmatrix},$$

$$\rho_2=\rho(\boldsymbol{x}_2)=(T\boldsymbol{x}_2,\ \boldsymbol{x}_2)=0,$$

再由
$$(T-\rho_2)\boldsymbol{y}_3=\boldsymbol{x}_2,$$

即
$$\begin{pmatrix}0 & \frac{1}{2}\\ \frac{1}{2} & 0\end{pmatrix}\begin{pmatrix}y_{13}\\ y_{23}\end{pmatrix}=\begin{pmatrix}0\\ 1\end{pmatrix},$$

得 $\boldsymbol{y}_3=(2,0)^T, \boldsymbol{x}_3=(1,0)^T.$

由此可知：$\rho_k=0$, $\boldsymbol{x}_{2k-1}=(1,0)^T$, $\boldsymbol{x}_{2k}=(0,1)^T$. 尽管 ρ_k 始终是 0, 但是 0 不是特征值，T 的特征值是 $\theta_1=\frac{1}{2}$, $\theta_2=-\frac{1}{2}$, 对应的特征向量 $\boldsymbol{y}_1=(1,1)^T/\sqrt{2}$, $\boldsymbol{y}_2=(1,-1)^T/\sqrt{2}$. 恰有 $\rho_k=\theta_1+\theta_2$, $\boldsymbol{x}_{2k-1}=(\boldsymbol{y}_1+\boldsymbol{y}_2)/\sqrt{2}$, $\boldsymbol{x}_{2k}=(\boldsymbol{y}_1-\boldsymbol{y}_2)/\sqrt{2}$. 这一现象反映了 B. N. Parlett 和 W. Kahan 在 1969 年给出的一个定理.

定理：任意取一个单位向量 $\boldsymbol{x}_1$，作为初始向量，Rayleigh 商迭代法所产生的序列 ρ_k、$\boldsymbol{x}_k$，下列两者必有一个发生：

1. ρ_k 收敛到一个特征值，$\boldsymbol{x}_k$ 收敛到对应的特征向量. 收敛速度是三次的.

2. ρ_k 收敛到一个数 $\bar{\rho}$, $\bar{\rho}$ 不是 T 的特征值，$\boldsymbol{x}_{2k}$ 收敛到向量 $\boldsymbol{x}_+$, $\boldsymbol{x}_{2k-1}$ 收敛到向量 $\boldsymbol{x}_-$, $\boldsymbol{x}_++\boldsymbol{x}_-$ 和 $\boldsymbol{x}_+-\boldsymbol{x}_-$ 是 T 的两个特征向量，此时收敛速度是线性的.

有关这个定理的证明，可参看[10, 第四章]，不在此介绍了.

上述 Parlett 和 Kahan 定理指出 Rayleigh 商迭代收敛到特征值时，收敛速度是三次的；对于带 Rayleigh 商位移的 QL 方法，如果 $\beta_1^{(k)}\to 0$，那么它的收敛速度也是三次的，有如下定理：

定理 3.12 设 T 不可约，它的特征值为 θ_1、θ_2、…、θ_n 由带 Rayleigh 商位移的 QL 方法，产生 $\{T^{(k)}\}$，如果 $\beta_1^{(k)}\to 0$，则 $\alpha_1^{(k)}\to\theta_i$，并且

$$|\beta_1^{(k+1)}|=|\beta_1^{(k)}|^3|\det(T_{3,n}^{(k)}-\sigma_k I)|(\omega_2^{(k)})^{1/2}/|\omega_1^{(k)}|, \tag{29}$$

这里

$$\omega_1^{(k)}=\det(T_{2,n}^{(k)}-\sigma_k I)^2+(\det(T_{3,n}^{(k)}-\sigma_k I)\beta_1)^2+\cdots$$
$$+(\beta_1^{(k)}\beta_2^{(k)}\cdots\beta_{n-1}^{(k)})^2,$$

有

$$\lim_{k\to\infty}\omega_1^{(k)}=\left(\prod_{\substack{l=1\\l\neq i}}^{n}(\theta_l-\theta_i)\right)^2\neq 0, \tag{30}$$

$$\omega_2^{(k)}=\det(T_{3,n}^{(k)}-\sigma_k I)^2+(\det(T_{4,n}^{(k)}-\sigma_k I)\beta_2)^2+\cdots$$
$$+(\beta_2^{(k)}\beta_3^{(k)}\cdots\beta_{n-1}^{(k)})^2$$

是一致有界量.

证明 $\beta_1^{(k)}\to 0$, 因此$\alpha_1^{(k)}=\sigma_k\to\theta_i$(参见[31]), θ_i是T的某个特征值。由此可知 $|\det(T_{l,n}^{(k)}-\sigma_k I)|$ 是一致有界的，另外 $|\beta_i^{(k)}|=|(T^{(k)}e_{i+1}, e_i)|\leqslant\|T^{(k)}\|=\|T\|$, $i=1, 2, \cdots, n-1$, 因此 $\omega_2^{(k)}$ 是一致有界量.

若 $T_{2,n}^{(k)}$ 的特征值为 $\theta_1^{(k)}, \theta_2^{(k)}, \cdots, \theta_{n-1}^{(k)}$, 由估计式 (9) 知 $\theta_1^{(k)}$、$\theta_2^{(k)}$、$\cdots$、$\theta_{n-1}^{(k)}$ 收敛到 θ_{i_1}、θ_{i_2}、$\cdots$、$\theta_{i_{n-1}}$,

$$\det(T_{2,n}^{(k)}-\sigma_k I)=\prod_{l=1}^{n-1}(\theta_l^{(k)}-\sigma_k),$$

故
$$\lim_{k\to\infty}\det(T_{2,n}^{(k)}-\sigma_k I)=\prod_{\substack{l=1\\l\neq i}}^{n}(\theta_l-\theta_i),$$

由此可得
$$\lim_{k\to\infty}\omega_1^{(k)}=\left(\prod_{\substack{l=1\\l\neq i}}^{n}(\theta_l-\theta_i)\right)^2,$$

因 T 不可约，于是 T 的特征值全不相同，故上式不等于0.

利用定理 3.10, 有

$$|\beta_1^{(k+1)}|=|\beta_1^{(k)}|\,|\det(T^{(k)}-\sigma_k I)|\,(\omega_2^{(k)})^{\frac{1}{2}}/|\omega_1^{(k)}|, \tag{31}$$

但

$$\det(T^{(k)}-\sigma_k I)=(\alpha_1^{(k)}-\sigma_k)\det(T_{2,n}^{(k)}-\sigma_k I)$$
$$-(\beta_1^{(k)})^2\det(T_{3,n}^{(k)}-\sigma_k I),$$

因为 $\alpha_1^{(k)}=\sigma_k$, 故

$$\det(T^{(k)}-\sigma_k I)=-(\beta_1^{(k)})^2\det(T_{3,n}^{(k)}-\sigma_k I),$$

将此代入(31)即得(29)式. 证毕.

由定理 3.12 可知 $\beta_1^{(k)}\to 0$ 时, 它的收敛速度至少是三次的. 但是满足什么样条件的对称三对角矩阵, 可以保证上述 $\beta_1^{(k)}\to 0$, 仍然是一个值得研究的问题.

§4 带 Wilkinson 位移的 QL 方法

鉴于带 Rayleigh 商位移的 QL 方法,偶而有 $\beta_1^{(k)}$ 不收敛于 0 的情况发生,目前用的较多的是带 Wilkinson 位移的QL方法. 这种位移取法在 Wilkinson 1965 出版的书[20]中已经提到. 1968 年 Wilkinson 在文[21]中证明: 这种取法的位移,能保证对任意对称三对角矩阵,$\beta_1^{(k)}$ 都收敛于 0,并证明收敛速度至少是二次的. 但是证明比较复杂,也可参看 C. L. Lawson 和 R. J. Hanson 的书[22, Appendix B].

1978 年 Hoffman 和 Parlett 给出一个很巧妙的关于 $\beta_1^{(k)}$ 收敛于 0 的证明, 同时也给出在一般情况下 $\hat{\beta}_1\sim 0(\beta_1^3\beta_2^2)$, 在特殊情况下 $\hat{\beta}_1\sim 0(\beta_1^2)$ 的收敛速度估计. 不过在证明中, 用到这样的命题: 对于方程

$$(T-\sigma I)p=e_1$$

的解 $p=(\pi_1, \pi_2, \pi_3, \cdots, \pi_n)^T$, 当 β_1 充分小时它的分量中模最大的是 π_1 和 π_2. 这个命题没有严格的证明. B. N. Parlett 在 1980 年出版的书[10, p. 155～156]中,在附加条件

$$\beta_2\to 0,\quad \beta_3\to 0,\quad \alpha_i\to\theta_i\quad (i=1、2、3), \tag{32}$$

证明了 $\quad |\hat{\beta}_1/\beta_1^3\beta_2|\to 1/|\theta_2-\theta_1|^3|\theta_3-\theta_1|\neq 0.$

这里我们证明 $\beta_1^{(k)}\to 0$, 是采用书[10]中的证明, 关于收敛速度的估计我们去掉了附加的假定 (32), 给出 $\hat{\beta}_1$ 与 β_1、

β_2、β_3、… 的正确关系式. 指出如果 $\lim_{k\to\infty}(\alpha_2^{(k)}-\sigma_k)\neq 0$, 则 $\hat{\beta}_1=o(\beta_1^3\beta_2^2)$. 任何情况下 $\hat{\beta}_1=o(|\alpha_1-\sigma|\beta_1\beta_2^2)$.

先介绍什么叫 Wilkinson 位移. 对于

$$T=\begin{pmatrix} \alpha_1 & \beta_1 & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1} \\ 0 & & & \beta_{n-1} & \alpha_n \end{pmatrix},$$

它的二阶顺序主子阵$\begin{pmatrix} \alpha_1 & \beta_1 \\ \beta_1 & \alpha_2 \end{pmatrix}$, 有两个实根

$$\omega_1=[\alpha_1+\alpha_2+\sqrt{(\alpha_1-\alpha_2)^2+4\beta_1^2}]/2,$$
$$\omega_2=[\alpha_1+\alpha_2-\sqrt{(\alpha_1-\alpha_2)^2+4\beta_1^2}]/2,$$

取其中靠近 α_1 的那个根, 作为位移值 σ, 带这样取法的位移, 称为带 Wilkinson 位移. 当然对于 $T^{(k)}$ 来说上述 α_1、α_2、β_1 和 σ 都要标以 k 的.

如果记 $\delta=(\alpha_2-\alpha_1)/2$, 那么

$$\omega_1=\frac{\alpha_1+\alpha_2}{2}+\sqrt{\delta^2+\beta_1^2}=\frac{\alpha_1+\alpha_2}{2}-\delta+\delta+\sqrt{\delta^2+\beta_1^2}$$
$$=\alpha_1+\delta+\sqrt{\delta^2+\beta_1^2}=\alpha_1-\beta_1^2/(\delta-\sqrt{\delta^2+\beta_1^2}),$$

即
$$\omega_1-\alpha_1=-\beta_1^2/(\delta-\sqrt{\delta^2+\beta_1^2}),$$

同样
$$\omega_2-\alpha_1=-\beta_1^2/(\delta+\sqrt{\delta^2+\beta_1^2}),$$

因此当 $\delta>0$ 时

$$|\omega_2-\alpha_1|<|\omega_1-\alpha_1|,$$

当 $\delta<0$ 时
$$|\omega_1-\alpha_1|<|\omega_2-\alpha_1|.$$

因此当 $\delta\geqslant 0$ 取 $\sigma=\omega_2$, 当 $\delta<0$ 时取 $\sigma=\omega_1$, 利用函数

$$\operatorname{sign}\delta=\begin{cases} 1, & \delta\geqslant 0; \\ -1, & \delta<0, \end{cases}$$

不管那种情况可取

$$\sigma=\alpha_1-\operatorname{sign}\delta(\beta_1^2/(|\delta|+\sqrt{\delta^2+\beta_1^2})). \tag{33}$$

因为 σ 是 $\begin{pmatrix}\alpha_1 & \beta_1\\ \beta_1 & \alpha_1\end{pmatrix}$ 的特征值，因此

$$(\alpha_1-\sigma)(\alpha_2-\sigma)=\beta_1^2,$$

从(33)可知

$$|\alpha_1-\sigma|\leqslant\beta_1,$$

等式只有当 $\delta=0$ 时才成立．同样

$$|\alpha_1-\sigma|\leqslant|\alpha_2-\sigma|,$$

等式只有当 $\delta=0$ 时成立．这两个关系，可以写成

$$\frac{|\alpha_1-\sigma|}{|\beta_1|}=\frac{|\beta_1|}{|\alpha_2-\sigma|}=\sqrt{\frac{|\alpha_1-\sigma|}{|\alpha_2-\sigma|}}\leqslant 1,$$

最后的 $\leqslant$ 号中的等号只有当 $\delta=0$ 时才成立．

现在来考察，带 Wilkinson 位移时，$\det(T-\sigma I)$ 是什么？从

$$\det(T-\sigma I)=(\alpha_1-\sigma)\det(T_{2,n}-\sigma I)-\beta_1^2\det(T_{3,n}-\sigma I),$$

而

$$\begin{aligned}\det(T_{2,n}-\sigma I)=&(\alpha_2-\sigma)\det(T_{3,n}-\sigma I)\\&-\beta_2^2\det(T_{4,n}-\sigma I),\end{aligned}$$

于是

$$\begin{aligned}\det(T-\sigma I)=&(\alpha_1-\sigma)[(\alpha_2-\sigma)\det(T_{3,n}-\sigma I)\\&-\beta_2^2\det(T_{4,n}-\sigma I)]-\beta_1^2\det(T_{3,n}-\sigma I)\\=&[(\alpha_1-\sigma)(\alpha_2-\sigma)-\beta_1^2]\det(T_{3,n}-\sigma I)\\&-(\alpha_1-\sigma)\beta_2^2\det(T_{4,n}-\sigma I),\end{aligned}$$

因为 $(\alpha_1-\sigma)(\alpha_2-\sigma)-\beta_1^2=0$，故

$$\det(T-\sigma I)=-(\alpha_1-\sigma)\beta_2^2\det(T_{4,n}-\sigma I). \tag{34}$$

定理 3.13 设对任意对称不可约三对角阵 T，使用带 Wilkinson 位移 $\{\sigma_k\}$ 的 QL 方法，获矩阵序列 $\{T^{(k)}\}$，

$$T^{(k)}=\begin{pmatrix} \alpha_1^{(k)} & \beta_1^{(k)} & & & 0 \\ \beta_1^{(k)} & \alpha_2^{(k)} & \beta_2^{(k)} & & \\ & \beta_2^{(k)} & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{n-1}^{(k)} \\ 0 & & & \beta_{n-1}^{(k)} & \alpha_n^{(k)} \end{pmatrix},$$

则 $\beta_1^{(k)}\to 0$.

证明 若某个 σ_k 是 T 的特征值，则 $\beta_1^{(k+1)}=0$ 且 $\beta_1^{(l)}=0$, $l>k+1$，因此可设 σ_k 不是 T 的特征值．为了简明起见，在证明中省略序列指标 k，以 $\hat{\alpha}_i$、$\hat{\beta}_i$ 表示 $T^{(k+1)}$ 的元素，以 α_i、β_i 表示 $T^{(k)}$ 的元素．

利用 § 2 定理 3.9，对一切位移成立

$$l_{11}^2=\det(T-\sigma I)^2/\omega_1,$$

$$\omega_1=\det(T_{2,n}-\sigma I)^2+(\det(T_{3,n}-\sigma I)\beta_1)^2$$
$$+(\det(T_{4,n}-\sigma I)\beta_1\beta_2)^2+\cdots+(\beta_1\beta_2\cdots\beta_{n-1})^2,$$

$$(T-\sigma I)q_1=l_{11}e_1,$$

$$q_{11}=l_{11}\det(T_{2,n}-\sigma I)/\det(T-\sigma I),$$

$$q_{21}=-l_{11}\beta_1\det(T_{3,n}-\sigma I)/\det(T-\sigma I),$$

满足
$$(\alpha_1-\sigma)q_{11}+\beta_1q_{21}=l_{11},$$

或者

$$(\alpha_1-\sigma)\frac{\det(T_{2,n}-\sigma I)}{\det(T-\sigma I)}+\beta_1\left(-\beta_1\frac{\det(T_{3,n}-\sigma I)}{\det(T-\sigma I)}\right)=1,$$

因此点$\left(\dfrac{\det(T_{2,n}-\sigma I)}{\det(T-\sigma I)},\ -\beta_1\dfrac{\det(T_{3,n}-\sigma I)}{\det(T-\sigma I)}\right)$在直线

$$(\alpha_1-\sigma)x+\beta_1y=1$$

上．而该直线离原点的最短距离为 $1/\sqrt{(\alpha_1-\sigma)^2+\beta_1^2}$，故

$$\left(\frac{\det(T_{2,n}-\sigma I)}{\det(T-\sigma I)}\right)^2+\left(\beta_1\frac{\det(T_{3,n}-\sigma I)}{\det(T-\sigma I)}\right)^2$$

$$\geqslant\frac{1}{(\alpha_1-\sigma)^2+\beta_1^2},$$

$$l_{11}^2\leqslant 1\Big/\left[\left(\frac{\det(T_{2,n}-\sigma I)}{\det(T-\sigma I)}\right)^2+\left(\beta_1\frac{\det(T_{3,n}-\sigma I)}{\det(T-\sigma I)}\right)^2\right]$$

$$\leqslant 1/((\alpha_1-\sigma)^2+\beta_1^2)^{-1}=(\alpha_1-\sigma)^2+\beta_1^2, \tag{35}$$

这是对一切位移σ都成立的估计式. 现在假如取的σ是Wilkinson位移,于是

$$(\alpha_1-\sigma)^2\leqslant\beta_1^2,$$

故

$$l_{11}^2\leqslant 2\beta_1^2. \tag{36}$$

再从

$$l_{11}^2\leqslant 1\Big/\left[\left(\frac{\det(T_{2,n}-\sigma I)}{\det(T-\sigma I)}\right)^2\right.$$

$$\left.+\left(\frac{\beta_1\det(T_{3,n}-\sigma)}{\det(T-\sigma)}\right)^2+\left(\beta_1\beta_2\frac{\det(T_{4,n}-\sigma)}{\det(T-\sigma)}\right)^2\right].$$

利用

$$\det(T-\sigma I)=-(\alpha_1-\sigma)\beta_2^2\det(T_{4,n}-\sigma I),$$

故

$$l_{11}^2\leqslant 1\Big/\left[\frac{1}{(\alpha_1-\sigma)^2+\beta_1^2}+\left(\frac{\beta_1\beta_2}{(\alpha_1-\sigma)\beta_2^2}\right)^2\right]$$

$$\leqslant 1\Big/\left(\frac{1}{2\beta_1^2}+\left(\frac{\beta_1}{(\alpha_1-\sigma)\beta_2}\right)^2\right)$$

$$\leqslant 1\Big/\left(\frac{1}{2\beta_1^2}+\left(\frac{1}{\beta_2}\right)^2\right)=\frac{2\beta_1^2\beta_2^2}{\beta_2^2+2\beta_1^2}$$

$$=\frac{(\sqrt{2}\,|\beta_1|)\,|\beta_2|\,(\sqrt{2}\,|\beta_1|)\,|\beta_2|}{(\sqrt{2}\,\beta_1)^2+\beta_2^2}\leqslant\frac{1}{2}\sqrt{2}\,|\beta_1|\,|\beta_2|. \tag{37}$$

下面先证 $\beta_1^2\beta_2\to 0$，因为

$$\hat{T}=Q^*TQ,$$

是从一个对称三对角阵正交相似变换到另一对称三对角阵，故可应用第2章§3的极值性质

$$|\hat{\beta}_1\hat{\beta}_2|=\min_{\psi_2\in mp_2}\|\psi_2(T)q_1\|,$$

特别取 $\psi_2(\lambda)=(\lambda-\alpha_1)(\lambda-\sigma)$，于是

$$|\hat{\beta}_1\hat{\beta}_2|\leqslant\|(T-\alpha_1)(T-\sigma)q_1\|$$

$$=\|(T-\alpha_1)l_{11}e_1\|=l_{11}|\beta_1|,$$

利用定理3.8的推论 $|\hat{\beta}_1|\leqslant l_{11}$，故

$$|\hat{\beta}_1^2\hat{\beta}_2|\leqslant l_{11}^2|\beta_1|,$$

将(37)代入上式，即得

$$|\hat{\beta}_1^2\hat{\beta}_2|\leqslant\frac{\sqrt{2}}{2}|\beta_1^2\beta_2|,$$

而 $\sqrt{2}/2=0.7071067\cdots<1$，这就证明了

$$\beta_1^2\beta_2\to 0.$$

另一方面

$$|\hat{\beta}_1|^3=|\hat{\beta}_1||\hat{\beta}_1|^2\leqslant l_{11}l_{11}^2$$

$$\leqslant\sqrt{2}|\beta_1|\frac{\sqrt{2}}{2}|\beta_1||\beta_2|=|\beta_1^2\beta_2|,$$

故

$$|\beta_1^3|\to 0,$$

从而证明了 $\beta_1\to 0$. 证毕.

下面的定理给出收敛速度的估计.

定理3.14 对任意对称不可约三对角阵 T，使用带 Wilkinson 位移的 QL 方法，则 $T^{(k+1)}$ 的 $\hat{\beta}_1$ 与 $T^{(k)}$ 的 β_1 之间有下述关系：

$$|\hat{\beta}_1|=|\alpha_1-\sigma_k||\beta_1|\beta_2^2|\det(T_{4,n}^{(k)}-\sigma_kI)|(\omega_2^{(k)})^{\frac{1}{2}}/|\omega_1^{(k)}|, \tag{38}$$

这里

$$\omega_1^{(k)} = (\det(T_{2,n}^{(k)} - \sigma_k I))^2 + (\det(T_{3,n}^{(k)} - \sigma_k I)\beta_1)^2$$
$$+ (\det(T_{4,n}^{(k)} - \sigma_k I)\beta_1\beta_2)^2 + \cdots + (\beta_1\beta_2\cdots\beta_{n-1})^2$$
$$\to \left(\prod_{\substack{l=1 \\ l \neq i}}^{n} (\theta_l - \theta_i) \right)^2,$$

其中 θ_i 是 α_1 的极限.

$$\omega_2^{(k)} = (\det(T_{3,n}^{(k)} - \sigma_k I))^2 + (\det(T_{4,n}^{(k)} - \sigma_k I)\beta_2)^2$$
$$+ (\det(T_{5,n}^{(k)} - \sigma_k I)\beta_2\beta_3)^2 + \cdots + (\beta_2\beta_3\cdots\beta_{n-1})^2$$

是一致有界量.

特别当

$$\underline{\lim_{k\to\infty}} |\alpha_2^{(k)} - \sigma_k| \neq 0, \tag{39}$$

有关系式:

$$|\hat{\beta}_1| = |\beta_1|^3 \beta_2^2 |\det(T_{4,n}^{(k)} - \sigma_k I)| (\omega_2^{(k)})^{\frac{1}{2}} / (|\alpha_2 - \sigma_k| \omega_1^{(k)}). \tag{40}$$

证明 因为 $\beta_1 \to 0$, 故 $\alpha_1 - \sigma_k \to 0$, 因此 α_1 的极限存在, 且是 T 的特征值 θ_i, 参见[31].

利用定理 3.10 有

$$\hat{\beta}_1^2 = \beta_1^2 \det(T^{(k)} - \sigma_k I)^2 \omega_2^{(k)} / (\omega_1^{(k)})^2,$$

类似于定理 3.12 对于带 Rayleigh 位移的讨论，利用 $\beta_1 \to 0$ 有

$$\omega_1^{(k)} \to \prod_{\substack{l=1 \\ l \neq i}}^{n} (\theta_l - \theta_i)^2$$

和 $\omega_2^{(k)}$ 是一致有界量.

再利用 σ_k 是 Wilkinson 位移, 故成立

$$\det(T^{(k)} - \sigma_k I) = -(\alpha_1 - \sigma_k)\beta_2^2 \det(T_{4,n}^{(k)} - \sigma_k I),$$

于是 $\hat{\beta}_1^2 = (\alpha_1 - \sigma_k)^2 \beta_1^2 \beta_2^4 \det(T_{4,n}^{(k)} - \sigma_k I)^2 \omega_2^{(k)} / (\omega_1^{(k)})^2,$

即(38)式成立.

特别当 $\lim\limits_{k\to\infty}|\alpha_2^{(k)}-\sigma_k|\neq 0$，利用 $\alpha_1-\sigma_k=\beta_1^2/(\alpha_2-\sigma_k)$ 故得

$$|\hat{\beta}_1|=|\beta_1|^3\beta_2^2|\det(T_{4,n}^{(k)}-\sigma_k I)(\omega_2^{(k)})^{\frac{1}{2}}/(|\alpha_2-\sigma_k|\omega_1^{(k)}),$$

即(40)式成立．证毕．

从(38)式利用 $|\alpha_1-\sigma_k|\leqslant\beta_1$ 可以知道，β_1 的收敛速度至少是二次的．当条件(39)成立时，从(40)式可知收敛速度至少是三次的．因为 $\sigma_k\to\theta_i$，如果 $\beta_2\to 0$，那么 $\alpha_2\to\theta_j\neq\theta_i$，此时(39)一定成立，并且 $|\hat{\beta}_1|=O(|\beta_1|^3\beta_2^2)$ 成立．于是可知要 $|\hat{\beta}_1|=O(|\beta_1|^3\beta_2^2)$ 用不着再附加条件 $\beta_3\to 0$．这一点已说明定理 3.14 要比[10]中结果更强一些．

关于其他一些位移下，QL 方法的收敛性和收敛率的讨论，可参看[31]和[32]．在[32]中提出了一种称为 RW 位移的新的位移，它也保证对任意不可约对称三对角矩阵 T 情况下的 $\beta_1\to 0$，并且在任何情况下 β_1 收敛率至少是三次的．

第 4 章

解特征值问题的 Lanczos 算法

Lanczos 算法求对称矩阵的特征值是 1950 年提出的[23]，取定一个单位向量 $\boldsymbol{q}$，通过第 2 章所介绍 Lanczos 过程(7)构造一组正交化序列 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$、…、$\boldsymbol{q}_n$，$Q=[\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_n]$，则 $Q^*AQ=T$ 是一个对称三对角矩阵. 这样矩阵 A 的特征值问题，化成矩阵 T 的特征值问题，后者自然比较简单多了. 当然 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$ 计算到 $\boldsymbol{q}_j(j<n)$，$Q_j=[\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j]$ 得到 $j\times j$ 对称三对角阵 $T_j=Q_j^*AQ_j$，可以求 T_j 的特征值，作为 A 的特征值近似值，也是把求 A 的特征值问题化成简单的求 T_j 的特征值问题了. 这就是 Lanczos 算法求特征值的基本思想.

但是由于舍入误差的影响，在求 $\boldsymbol{q}_i$ 的过程中，所求到的 $\boldsymbol{q}_i$ 很快就失去了正交性，因此 Lanczos 算法 50 年代、60 年代被认为是不稳定的，很少被人使用. 直到 1971 年 C. C. Paige 在他的著名博士论文[13]中，通过舍入误差的分析，发现失去正交性，恰与近似特征值的精度提高相关. 重新肯定 Lanczos 算法是一种求大稀疏矩阵，两端部的特征值，即最大的部分和最小的部分特征值的一种有效方法. 自那以后 Lanczos 算法越来越被重视，这是因为它具有很大的优点:

1. 每次迭代用到原来矩阵 A 的只是 A 与向量的乘法. 因此可以充分利用 A 的稀疏性.

2. 在一次迭代中要进行的计算，都是向量间的运算，有利于实现平行运算.

3. 在一次迭代中，只用到二个向量，其余已算得的资料

可以放在外存贮中.

4. 对计算的结果,可以有误差的估计.

因为有上述这些优点，因此目前对于非常高阶的稀疏对称矩阵的特征值问题,都用 Lanczos 算法来解.

§1 近似不变子空间

设 A 是 $n\times n$ 实对称矩阵,它的 n 个特征值，由小到大排成 $\lambda_1, \lambda_2, \cdots, \lambda_n$,对应的单位正交特征向量为 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n$.

对任意非零 $x\in R^n$, 定义 Rayleigh 商为

$$R(\boldsymbol{x})=(A\boldsymbol{x}, \boldsymbol{x})/(\boldsymbol{x}, \boldsymbol{x})$$

有如下极大极小原理

定理4.1 记 V_l 是一个任意 l 维子空间,$l\leqslant n$, 则

$$\lambda_l=\min_{V_l}\max_{x\in V_l}R(\boldsymbol{x}), \tag{1}$$

$$\lambda_{n-l+1}=\max_{V_l}\min_{x\in V_l}R(\boldsymbol{x}). \tag{2}$$

证明 记 $n-l+1$ 维空间 $E_{l-1}^{\perp}=\{\boldsymbol{y}_l, \boldsymbol{y}_{l+1}, \cdots, \boldsymbol{y}_n\}$, 于是任意 l 维空间 V_l 与 $E_{l-1}^{\perp}$ 的交不空. 取非零向量 $\boldsymbol{y}\in V_l\cap E_{l-1}^{\perp}$ 有

$$\max_{x\in V_l}R(x)\geqslant R(\boldsymbol{y})\geqslant\lambda_l,$$

故

$$\min_{V_l}\max_{x\in V_l}R(\boldsymbol{x})\geqslant\lambda_l.$$

另一方面, 取一个 l 维子空间 $\hat{V}_l=\{\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_l\}$,

$$\max_{x\in\hat{V}_l}R(\boldsymbol{x})=\lambda_l\geqslant\min_{V_l}\max_{x\in V_l}R(\boldsymbol{x}),$$

从而(1)成立. 同样可证(2)成立. 证毕.

定理4.2 设 E 也是 $n\times n$ 对称矩阵,它的特征值由小到大排成 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$, 如果将 $A+E$ 的 n 个特征值由小到大

排成 $\mu_1, \mu_2, \cdots, \mu_n$, 则

$$\mu_l \leqslant \lambda_l + \varepsilon_n, \tag{3}$$

$$\lambda_l \leqslant \mu_l - \varepsilon_1. \tag{4}$$

证明 由定理 4.1

$$\mu_l = \min_{V_l} \max_{x \in V_l} ((A+E)\boldsymbol{x}, \boldsymbol{x})/(\boldsymbol{x}, \boldsymbol{x})$$

$$= \min_{V_l} \max_{x \in V_l} \left[\frac{(A\boldsymbol{x}, \boldsymbol{x})}{(\boldsymbol{x}, \boldsymbol{x})} + \frac{(E\boldsymbol{x}, \boldsymbol{x})}{(\boldsymbol{x}, \boldsymbol{x})} \right]$$

$$\leqslant \max_{x \in \hat{V}_l} \left[\frac{(A\boldsymbol{x}, \boldsymbol{x})}{(\boldsymbol{x}, \boldsymbol{x})} + \frac{(E\boldsymbol{x}, \boldsymbol{x})}{(\boldsymbol{x}, \boldsymbol{x})} \right],$$

这里 $\hat{V}_l = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_l\}$. 于是

$$\mu_l \leqslant \max_{x \in \hat{V}_l} \frac{(A\boldsymbol{x}, \boldsymbol{x})}{(\boldsymbol{x}, \boldsymbol{x})} + \max_{x \in \hat{V}_l} \frac{(E\boldsymbol{x}, \boldsymbol{x})}{(\boldsymbol{x}, \boldsymbol{x})} \leqslant \lambda_l + \varepsilon_n.$$

由 $A = A+E-E$, 而 $-E$ 的最大特征值为 $-\varepsilon_1$, 利用上述结果就得

$$\lambda_l \leqslant \mu_l - \varepsilon_1. \quad \text{证毕.}$$

推论

$$|\lambda_l - \mu_l| \leqslant \|E\|. \tag{5}$$

设 $Q = [\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j]$ 是 j 个 n 维、线性无关的列向量组成的矩阵, $j<n$, 若存在 $j \times j$ 矩阵 T_j 使

$$AQ - QT_j = 0,$$

那么 $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j\}$ 是 A 的不变子空间. 此时 T_j 的特征值, 即为 A 的特征值, 若 $\boldsymbol{s}$ 是 T_j 的特征向量, 则 $Q\boldsymbol{s}$ 是 A 的特征向量. 如果对于 Q, 有 T_j 使得

$$AQ - QT_j = R,$$

$\|R\|$ 很小, 那么可以设想 $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_n\}$ 是 A 的近似不变子空间. 下面我们要讨论此时 T_j 的特征值、特征向量与 A 的特征值、特征向量之间的关系.

对于给定的 Q，任意取一个 $j\times j$ 矩阵 T_j，都有一个 $R(T_j)$，使得

$$AQ-QT_j=R(T_j),$$

首先有一个问题是：选取怎样的 T_j，使 $\|R(T_j)\|$ 最小。

对于 $j=1$ 的情况，T_j 是个常数 α，考虑

$$A\boldsymbol{q}_1-\boldsymbol{q}_1\alpha=\boldsymbol{r},$$

问 α 取什么值时，使 $\|\boldsymbol{r}\|=\|A\boldsymbol{q}_1-\boldsymbol{q}_1\alpha\|$ 最小。这问题可以看成向量 $A\boldsymbol{q}_1$，在 $\boldsymbol{q}_1$ 所张成的子空间中找一个最佳逼近向量，利用第1章的定理1.1，我们知道当 α 使

$$(A\boldsymbol{q}_1-\alpha\boldsymbol{q}_1,\ \boldsymbol{q}_1)=0$$

时，$\|\boldsymbol{r}\|$ 最小，也即取

$$\alpha^0=(A\boldsymbol{q}_1,\ \boldsymbol{q}_1)/(\boldsymbol{q}_1,\ \boldsymbol{q}_1), \tag{6}$$

有 $$\|A\boldsymbol{q}_1-\alpha^0\boldsymbol{q}_1\|\leqslant\|A\boldsymbol{q}_1-\alpha\boldsymbol{q}_1\|,\quad \forall\alpha\in R,$$

注意(6)式恰为 $\boldsymbol{q}_1$ 的 Rayleigh 商，因此可以说：对于 $A\boldsymbol{q}_1-\boldsymbol{q}_1\alpha$ 当 α 取成 $\boldsymbol{q}_1$ 的 Rayleigh 商时，使 $\|A\boldsymbol{q}_1-\boldsymbol{q}_1\alpha\|$ 最小。

对于任意的 α，它跟 A 的特征值之间成立如下定理中所述的估计式。

定理4.3 对于 α，可以找到一个 A 的特征值 λ，使得

$$|\lambda-\alpha|\leqslant\|r\|/\|\boldsymbol{q}_1\|. \tag{7}$$

证明 将 $\boldsymbol{q}_1$ 按 A 的特征向量展开

$$\boldsymbol{q}_1=\sum_{i=1}^{n}\beta_i\boldsymbol{y}_i,$$

于是 $$\boldsymbol{r}=\sum_{i=1}^{n}\beta_i(\lambda_i-\alpha)\boldsymbol{y}_i,$$

$$\begin{aligned}\|\boldsymbol{r}\|^2&=\sum_{i=1}^{n}\beta_i^2(\lambda_i-\alpha)^2\\&\geqslant\min_i\ (\lambda_i-\alpha)^2\sum_{i=1}^{n}\beta_i^2=\min_i\ (\lambda_i-\alpha)^2\|\boldsymbol{q}_1\|^2,\end{aligned}$$

故 $$\min_i (\lambda_i-\alpha)^2 \leqslant \|\boldsymbol{r}\|^2/\|\boldsymbol{q}_1\|^2.$$ 证毕.

在定理证明中使$(\lambda_i-\alpha)^2$达到极小的那个特征值λ，如果对应的特征向量不止一个，那末$\boldsymbol{q}_1$的分解式，可以写成

$$\boldsymbol{q}_1=\sum_{\lambda_i=\lambda}\beta_i\boldsymbol{y}_i+\sum_{\lambda_i\neq\lambda}\beta_i\boldsymbol{y}_i,$$

将$\sum\limits_{\lambda_i=\lambda}\beta_i\boldsymbol{y}_i$重新写成$\beta\boldsymbol{y}$，其中$\boldsymbol{y}$是单位向量，显然也是对应特征值$\lambda$的特征向量．于是

$$\boldsymbol{q}_1=\beta\boldsymbol{y}+\sum_{\lambda_i\neq\lambda}\beta_i\boldsymbol{y}_i,$$

假定$\|\boldsymbol{q}_1\|=1$，可得$\beta^2+\sum\limits_{\lambda_i\neq\lambda}\beta_i^2=1$.

定理 4.4 如果$d=\min\limits_{\lambda_i\neq\lambda}|\lambda_i-\alpha|$，则

$$\|\boldsymbol{q}_1-\beta\boldsymbol{y}\|\leqslant\|\boldsymbol{r}\|/d.$$

证明 从$A\boldsymbol{q}_1-\alpha\boldsymbol{q}_1=\boldsymbol{r}$有

$$\beta(\lambda-\alpha)\boldsymbol{y}+\sum_{\lambda_i\neq\lambda}\beta_i(\lambda_i-\alpha)\boldsymbol{y}_i=\boldsymbol{r},$$

$$\|\boldsymbol{r}\|^2\geqslant\sum_{\lambda_i\neq\lambda}\beta_i^2(\lambda_i-\alpha)^2\geqslant d^2\sum_{\lambda_i\neq\lambda}\beta_i^2,$$

故 $$\sum_{\lambda_i\neq\lambda}\beta_i^2\leqslant\|\boldsymbol{r}\|^2/d^2,$$

但 $$\boldsymbol{q}_1-\beta\boldsymbol{y}=\sum_{\lambda_i\neq\lambda}\beta_i\boldsymbol{y}_i,$$

故 $$\|\boldsymbol{q}_1-\beta\boldsymbol{y}\|^2\leqslant\sum_{\lambda_i\neq\lambda}\beta_i^2\leqslant\|\boldsymbol{r}\|^2/d^2.$$ 证毕.

推论 若$\boldsymbol{q}_1$与$\boldsymbol{y}$的交角$\angle(\boldsymbol{q}_1,\ \boldsymbol{y})=\theta$，则

$$|\sin\theta|\leqslant\|\boldsymbol{r}\|/d. \tag{8}$$

证明 因为$(\boldsymbol{q}_1,\ \boldsymbol{y})=\beta=\cos\theta$，故$\sin^2\theta=1-\beta^2$,

$$\sin^2\theta=\sum_{\lambda_i\neq\lambda}\beta_i^2,$$

从而 $$|\sin\theta|\leqslant\|\boldsymbol{r}\|/d.$$

上述定理描写了$\boldsymbol{q}_1$逼近A的特征向量的状况．特别当α取成$\boldsymbol{q}_1$的 Rayleigh 商$\rho(\boldsymbol{q}_1)$时，有

定理 4.5 设 $\|\boldsymbol{q}_1\|=1$, $\rho=\rho(\boldsymbol{q}_1)$, $\boldsymbol{r}=A\boldsymbol{q}_1-\rho\boldsymbol{q}_1$, $d=\min\limits_{\lambda_i\neq\lambda}|\lambda_i-\rho|$, λ 是与 ρ 最接近的特征值,则

$$|\lambda-\rho|\leqslant\|r\|^2/d. \tag{9}$$

证明 利用分解式

$$\boldsymbol{q}_1=\beta\boldsymbol{y}+\sum_{\lambda_i\neq\lambda}\beta_i\boldsymbol{y},$$

$$\boldsymbol{r}=\beta(\lambda-\rho)\boldsymbol{y}+\sum_{\lambda_i\neq\lambda}\beta_i(\lambda_i-\rho)\boldsymbol{y}_i,$$

有
$$\|\boldsymbol{r}\|^2=\beta^2(\lambda-\rho)^2+\sum_{\lambda_i\neq\lambda}\beta_i^2(\lambda_i-\rho)^2.$$

利用 $\boldsymbol{q}_1^*\boldsymbol{r}=0$, 故

$$\beta^2(\lambda-\rho)+\sum_{\lambda_i\neq\lambda}\beta_i^2(\lambda_i-\rho)=0,$$

或
$$\beta^2(\lambda-\rho)=-\sum_{\lambda_i\neq\lambda}\beta_i^2(\lambda_i-\rho),$$

代入 $\|\boldsymbol{r}\|^2$ 的表示式中

$$\begin{aligned}\|\boldsymbol{r}\|^2&=-(\lambda-\rho)\sum_{\lambda_i\neq\lambda}\beta_i^2(\lambda_i-\rho)+\sum_{\lambda_i\neq\lambda}\beta_i^2(\lambda_i-\rho)^2\\&=\sum_{\lambda_i\neq\lambda}\beta_i^2(\lambda_i-\lambda)(\lambda_i-\rho),\end{aligned}$$

又从
$$\rho-\lambda=\boldsymbol{q}_1^*(A-\lambda)\boldsymbol{q}_1=\sum_{\lambda_i\neq\lambda}\beta_i^2(\lambda_i-\lambda),$$

故

$$|\rho-\lambda|\leqslant\sum_{\lambda_i\neq\lambda}\beta_i^2|\lambda_i-\lambda|\leqslant\frac{1}{d}\sum_{\lambda_i\neq\lambda}\beta_i^2|\lambda_i-\lambda|\,|\lambda_i-\rho|. \tag{10}$$

因为 λ 是 ρ 最接近的特征值,故有

$$|\lambda_i-\lambda|\,|\lambda_i-\rho|=(\lambda_i-\lambda)(\lambda_i-\rho),$$

将此代入(10)即得

$$|\rho-\lambda|\leqslant\frac{1}{d}\sum_{\lambda_i\neq\lambda}\beta_i^2(\lambda_i-\lambda)(\lambda_i-\rho)=\|\boldsymbol{r}\|^2/d.$$ 证毕.

推论 若 λ 是 λ_1 或 λ_n, 则

$$|\tan\theta|\leqslant\|\boldsymbol{r}\|/d.$$

证明 由

$$\beta^2(\lambda-\rho) = -\sum_{\lambda_i\neq\lambda}\beta_i^2(\lambda_i-\rho),$$

因为所有 $\lambda_i\neq\lambda$ 的 $\lambda_i-\rho$ 都是同号的，因此

$$|\lambda-\rho|\beta^2 = \sum_{\lambda_i\neq\lambda}\beta_i^2|\lambda_i-\rho| \geqslant d\sum_{\lambda_i\neq\lambda}\beta_i^2,$$

$$\tan^2\theta = \sum_{\lambda_i\neq\lambda}\beta_i^2/\beta^2 \leqslant |\lambda-\rho|/d \leqslant \|\boldsymbol{r}\|^2/d^2.$$ 证毕.

这个定理和推论说明当 α 取成 Rayleigh 商，其估计有所加强.

现在来考虑，Q 有 j 个列，$j>1$ 的情况，设 $Q=[\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j]$，$\boldsymbol{q}_1$、$\boldsymbol{q}_2$、$\cdots$、$\boldsymbol{q}_j$ 是单位正交向量组，

$$AQ-QT_j=R,$$

讨论 T_j 的特征值与 A 的特征值的关系. 首先有

定理 4.6 对任意 T_j 的特征值 μ，可以找到一个 A 的特征值 λ，使得

$$|\lambda-\mu| \leqslant \|R\|.$$

证明 μ 是 T_j 的特征值，记 $\boldsymbol{s}$ 是它对应的单位特征向量，令 $\boldsymbol{x}=Q\boldsymbol{s}$，于是

$$\|\boldsymbol{x}\| = \|\boldsymbol{s}^*Q^*Q\boldsymbol{s}\|^{1/2}=1,$$

$$AQ\boldsymbol{s}-QT_j\boldsymbol{s}=R\boldsymbol{s},$$

$$A\boldsymbol{x}-\mu\boldsymbol{x}=R\boldsymbol{s}.$$

利用定理 4.3，有 A 的特征值 λ 使

$$|\lambda-\mu| \leqslant \|R\boldsymbol{s}\| \leqslant \|R\|.$$ 证毕.

这个定理的缺点是：对 T_j 的两个特征值 μ_1、μ_2，定理中所说的 A 的特征值 λ，可能是同一个. 即

$$|\lambda-\mu_1| \leqslant \|R\|,$$

$$|\lambda-\mu_2| \leqslant \|R\|.$$

人们希望对于 T_j 的 j 个特征值 μ_1、μ_2、$\cdots$、μ_j，对应找

到 A 的 j 个特征值 λ_{i_1}、λ_{i_2}、…、λ_{i_j}, i_1、i_2、…、i_j 是 j 个不同的数, 得到 $|\mu_k-\lambda_{i_k}|$ 的估计式.

这个任务在 1967 年, W. Kahan 的没有公开发表的著名论文: "Inclusion theorems for clusters of eigenvalues of Hermitian matrices" 中得到完成. 在那篇论文中, 他先证明了一个重要定理, 可称为矩阵的保范扩张定理. 这个定理第一次公开发表在 B. N. Parlett 的书[10]中, 在 [10] 中的证明, B. N. Parlett 作了适当的修改. 下面我们给出这个矩阵保范扩张定理, 是采用 W. Kahan 原来的证明.

定理 4.7 设 V 是 $n\times n$ Hermite 矩阵, C 是任意复数域上的 $m\times n$ 矩阵, 对于 $(n+m)\times n$ 矩阵

$$R=\begin{pmatrix} V \\ C \end{pmatrix}$$

存在 $m\times m$ Hermite 矩阵 W, 使得 $(n+m)\times(n+m)$ Hermite 矩阵

$$T=\begin{pmatrix} V & C^* \\ C & W \end{pmatrix}$$

有

$$\|T\|=\|R\|.$$

证明

$$R^*R=V^2+C^*C,$$

$$T^2=\begin{pmatrix} V^2+C^*C & VC^*+C^*W \\ CV+WC & CC^*+W^2 \end{pmatrix},$$

V^2+C^*C 是 T^2 的主子阵, 因此对任意 Hermite 矩阵 W, 都有

$$\|T^2\|\geqslant\|R\|^2, \quad 及 \quad \|T\|\geqslant\|R\|. \tag{11}$$

下面要找一个 W, 使得 $\|T\|$ 最小. 取任意数 $\rho>\|R\|$, 于是 ρ^2I-V^2 是一个 Hermite 正定矩阵, 当然有逆. 令 $D=C(\rho^2I-V^2)^{-1}$,

$$\rho^2 I - T^2 = \begin{pmatrix} \rho^2 I - V^2 - C^*C & -VC^* - C^*W \\ -CV - WC & \rho^2 I - CC^* - W^2 \end{pmatrix},$$

考察

$$\begin{pmatrix} I & 0 \\ DV & I \end{pmatrix} (\rho^2 I - T^2) \begin{pmatrix} I & VD^* \\ 0 & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ DV & I \end{pmatrix}$$

$$\times \begin{pmatrix} \rho^2 I - V^2 - C^*C & (\rho^2 I - V^2 - C^*C) VD^* - VC^* - C^*W \\ *\ *\ * & (-CV - WC) VD^* + \rho^2 I - CC^* - W^2 \end{pmatrix},$$

因此如果取 W 使

$$(\rho^2 I - V^2 - C^*C) VD^* - VC^* - C^*W$$
$$= -C^*CVD^* - C^*W = 0,$$

上述三个矩阵相乘的积，就是一个块对角阵．为此取

$$W(\rho) = -CVD^* = -DVC^*,$$

此时

$$\begin{pmatrix} I & 0 \\ DV & I \end{pmatrix} (\rho^2 I - T^2) \begin{pmatrix} I & VD^* \\ 0 & I \end{pmatrix} = \begin{pmatrix} \rho^2 I - R^*R & 0 \\ 0 & \rho^2 X \end{pmatrix},$$

$$X = I - C(\rho^2 I - V^2)^{-1} C^*.$$

另一方面

$$\begin{pmatrix} I & 0 \\ DV & I \end{pmatrix} (\rho^2 I - RR^*) \begin{pmatrix} I & VD^* \\ 0 & I \end{pmatrix}$$
$$= \begin{pmatrix} I & 0 \\ DV & I \end{pmatrix} \begin{pmatrix} \rho^2 I - V^2 & -VC^* \\ -CV & \rho^2 I - CC^* \end{pmatrix} \begin{pmatrix} I & DV \\ 0 & I \end{pmatrix}$$
$$= \begin{pmatrix} \rho^2 I - V^2 & 0 \\ 0 & \rho^2 X \end{pmatrix}.$$

因 RR^* 的最大特征值与 R^*R 的最大特征值相同，因此矩阵 $\rho^2 I - RR^*$ 也是正定矩阵，从而由惯性定理(或称 Sylvester 定理)[9, p. 335]知道 $\rho^2 X$ 也必须是正定矩阵，由此知道 $\rho^2 I - T^2$ 也必须是正定矩阵，即得

$$\rho>\|T\|,$$

再看 $$W(\rho)=-CV(\rho^2I-V^2)^{-1}C^*$$

是 ρ 的有理函数，当 $\rho\to\|R\|$ 时，如果 $\|W(\rho)\|$ 有界，则 $\rho=\|R\|$ 不是 $W(\rho)$ 的奇点(因为有理函数只可能有极点)，CC^*+W^2 也是 T^2 的主子阵，因此 T^2 的最大特征值大于 CC^*+W^2 的最大特征值，因为 CC^* 是非负定矩阵，故

$$\|W(\rho)\|\leqslant\|T\|<\rho$$

有界，故 $\rho\to\|R\|$，$W(\rho)$ 的极限 $W_+(\|R\|)$ 存在. 并且对于这个 $W(\|R\|)$，有

$$\|T\|\leqslant\|R\|,$$

联合(11)式，即有

$$\|T\|=\|R\|.$$ 证毕.

现在可以证明如下的 Kahan 定理.

定理 4.8 设 Q 是 j 个 n 维单位列正交向量组成，T_j 是 $j\times j$ 对称矩阵，

$$R=AQ-QT_j,$$

若 T_j 的 j 个特征值为 μ_1、μ_2、…、μ_j，则可以找到 j 个不同的自然数 i_1、i_2、…、i_j 使得

$$|\lambda_{i_k}-\mu_k|\leqslant\|R\|. \tag{12}$$

证明 将 Q 扩充成 $n\times n$ 正交矩阵 G，$G=[Q,\ Q_1]$，矩阵

$$G^*AG=\begin{pmatrix}Q^*AQ & Q^*AQ_1\\ Q_1^*AQ & Q_1^*AQ_1\end{pmatrix}$$

与 A 有相同的特征值.

$$\begin{aligned}G^*R&=G^*AGG^*Q-G^*QT_j\\&=\begin{pmatrix}Q^*AQ & Q^*AQ_1\\ Q_1^*AQ & Q_1^*AQ_1\end{pmatrix}\begin{pmatrix}I\\0\end{pmatrix}-\begin{pmatrix}I\\0\end{pmatrix}T_j\\&=\begin{pmatrix}Q^*AQ-T_j\\ Q_1^*AQ\end{pmatrix},\end{aligned}$$

因为 G^* 是正交阵，故 $\|G^*R\|=\|R\|$.

利用定理 4.7，扩充 G^*R 成

$$\begin{pmatrix} Q^*AQ-T_j & Q^*AQ_1 \\ Q_1^*AQ & W \end{pmatrix}=E,$$

有 W 存在，使得对称阵 E 的范数 $\|E\|=\|R\|$.

考察 $$G^*AG-E=\begin{pmatrix} T_j & 0 \\ 0 & Q_1^*AQ_1-W \end{pmatrix}.$$

利用定理 4.2，可知上述等式右端矩阵的第 l 个特征值，与 G^*AG 的第 l 个特征值之差的绝对值小于等于 $\|E\|$. 如果 T_j 的特征值 μ_1、μ_2、…、μ_j，在矩阵

$$\begin{pmatrix} T_j & 0 \\ 0 & Q_1^*AQ_1-W \end{pmatrix}$$

的特征值中，按由小到大次序排列的编号为 i_1、i_2、…、i_j，于是

$$|\lambda_{i_k}-\mu_k|\leqslant\|E\|=\|R\|.$$ 证毕.

这个定理加强了定理 4.6 的结果.

对于 $Q=[q_1, q_2, \cdots, q_j]$，$q_1$、$q_2$、…、$q_j$ 是 j 个单位正交列向量的情况，在等式

$$AQ-QT_j=R$$

中，也有一个使 $\|R\|$ 最小的最佳 T_j 问题. 下面定理回答了这个问题，见[10].

定理 4.9 对于给定的 j 个单位正交列向量组成的 $n\times j$ 矩阵 Q，对任意 $j\times j$ 矩阵 T_j，记

$$R(T_j)=AQ-QT_j,$$

若记 $j\times j$ 对称矩阵 $H=Q^*AQ$，则

$$\|R(H)\|\leqslant\|R(T_j)\|.$$

证明

$$R(T_j)^*R(T_j)=(AQ-QT_j)^*(AQ-QT_j)$$
$$=Q^*A^2Q-T_j^*Q^*AQ-Q^*AQT_j+T_j^*T_j$$
$$=Q^*A^2Q-H^2+(H-T_j)^*(H-T_j),$$

但 $$R(H)^*R(H)=Q^*A^2Q-H^2,$$

故 $R(T_j)^*R(T_j)=R(H)^*R(H)+(H-T_j)^*(H-T_j)$,

又因为$(H-T_j)^*(H-T_j)$是一个非负矩阵，于是

$$\|R(H)\|\leqslant\|R(T_j)\|.$$ 证毕.

W. Kahan 还考虑：矩阵 Q 的 j 个列不正交的情况下，T_j 的特征值，与 A 的特征值之间的关系. 在上面曾经指出过的那篇文章中他给出了如下定理，证明是按照[10]中所给出的方式.

定理4.10 设 Q 是任意 j 个线性无关的列向量构成的 $n\times j$ 矩阵，T_j 是任意 $j\times j$ 实对称矩阵，对于

$$R=AQ-QT_j,$$

若 Q 的最小奇异值为 σ_1，T_j 的特征值按由小到大次序排列是 μ_1、μ_2、…、μ_j，则存在 i_1、i_2、…、i_j，j 个不同自然数，使得

$$|\lambda_{i_k}-\mu_k|\leqslant\sqrt{2}\,\|R\|/\sigma_1. \tag{13}$$

证明 将 Q 进行奇异值分解，见[9, p. 468～470]. 存在 $n\times n$ 正交阵 P_1 和 $j\times j$ 正交阵 P_2，使得

$$P_1QP_2=\begin{pmatrix}J\\0\end{pmatrix},$$

J 为 $j\times j$ 对角阵 $\mathrm{diag}(\sigma_j,\ \sigma_{j-1},\ \cdots,\ \sigma_1)$. 对于

$$R=AQ-QT_j,$$

有 $$P_1RP_2=P_1AP_1^*P_1QP_2-P_1QP_2P_2^*T_jP_2.$$

记 $\hat{R}=P_1RP_2$, $\hat{A}=P_1AP_1^*$, $\hat{T}_j=P_2^*T_jP_2$, 有

$$\|\hat{R}\|=\|R\|,\quad \hat{A}\text{ 相似于 }A,\quad \hat{T}_j\text{ 相似于 }T_j.$$

$$\hat{R}=\hat{A}\begin{pmatrix}J\\0\end{pmatrix}-\begin{pmatrix}J\\0\end{pmatrix}\hat{T}_j,$$

若 $$\hat{A}=\begin{pmatrix}H & B^*\\ B & U\end{pmatrix}=\begin{pmatrix}\hat{T}_j & 0\\ 0 & U-X\end{pmatrix}+\begin{pmatrix}H-\hat{T}_j & B^*\\ B & X\end{pmatrix},$$

于是从定理 4.2 可知存在 i_1、i_2、$\cdots$、i_j 使得

$$|\lambda_{i_k}-\mu_k|\leqslant\left\|\begin{pmatrix}H-\hat{T}_j & B^*\\ B & X\end{pmatrix}\right\|,$$

利用定理 4.7，有 X 使

$$\left\|\begin{pmatrix}H-\hat{T}_j & B^*\\ B & X\end{pmatrix}\right\|=\left\|\begin{pmatrix}H-\hat{T}_j\\ B\end{pmatrix}\right\|,\tag{14}$$

下面来估计上式右端，易知

$$\hat{R}=\begin{pmatrix}HJ-J\hat{T}_j\\ BJ\end{pmatrix},$$

因为 Q 的列线性无关，因此 $\sigma_1>0$，故 J^{-1} 存在，

$$\|J^{-1}\|=1/\sigma_1,$$

$$\|B\|=\|BJJ^{-1}\|\leqslant\|BJ\|/\sigma_1,$$

而 $$HJ-J\hat{T}_j=(H-\hat{T}_j)J+(\hat{T}_jJ-J\hat{T}_j),$$

上式右端第二项是一个斜对称矩阵，对任意向量 $\boldsymbol{x}$，有

$$\boldsymbol{x}^*(\hat{T}_jJ-J\hat{T}_j)\boldsymbol{x}=0,$$

若对称矩阵 $H-\hat{T}_j$ 的模最大的特征值为 λ，对应的单位特征向量为 y，于是

$$y^*(HJ-J\hat{T}_j)y=\lambda y^*Jy,$$

即得 $$\|HJ-J\hat{T}_j\|\geqslant|\lambda|\sigma_1=\|H-\hat{T}_j\|\sigma_1.$$

再回到(14)式，记

$$M=\begin{pmatrix}H-\hat{T}_j\\ B\end{pmatrix},$$

$$M^*M=(H-\hat{T}_j,\ B^*)\begin{pmatrix}H-\hat{T}_j\\ B\end{pmatrix}=(H-\hat{T}_j)^2+B^*B,$$

所以

$$\|M^*M\|=\|(H-\hat{T}_j)^2+B^*B\|\leqslant\|H-\hat{T}_j\|^2+\|B\|^2$$
$$\leqslant(\|HJ-J\hat{T}_j\|^2+\|BJ\|^2)/\sigma_1^2,$$
$$\hat{R}^*\hat{R}=(HJ-J\hat{T}_j)^*(HJ-J\hat{T}_j)+(BJ)^*(BJ),$$

右端两项都是非负矩阵，故

$$\|\hat{R}\|\geqslant\|HJ-J\hat{T}_j\|,$$
$$\|\hat{R}\|\geqslant\|BJ\|,$$

从而得到 $\|M^*M\|\leqslant 2\|\hat{R}\|^2/\sigma_1^2$. 证毕.

当 Q 的各列单位正交时，按本定理结论，此时 $\sigma_1=1$，从(13)得

$$|\lambda_{i_k}-\mu_k|\leqslant\sqrt{2}\,\|R\|,$$

与定理 4.8 比较，多了一个 $\sqrt{2}$ 因子，因此可以设想 $\sqrt{2}$ 可以改成 1，但这个问题至今没有解决.

§2 Lanczos 算 法

给定一个单位向量 $\boldsymbol{q}_1$，构造序列

$$\boldsymbol{q}_1,\ A\boldsymbol{q}_1,\ A^2\boldsymbol{q}_1,\ \cdots,\ A^j\boldsymbol{q}_1, \tag{15}$$

$\{\boldsymbol{q}_1,\ A\boldsymbol{q}_1,\ \cdots,\ A^j\boldsymbol{q}_1\}$ 称为 Крылов 子空间. 在求线性方程组解时，系数矩阵 A 是对称正定的情况，共轭斜量法就是利用 Крылов 空间提供讯息来构造近似解的. 对于对称矩阵 A 的特征值问题. 很多求特征值的方法，也是利用 Крылов 子空间提供的讯息. 例如: 1931 年 Крылов 利用它来计算 $\boldsymbol{q}_1$ 的最小多项式. 这一方法在 n 很小时是可以用的，当 n 较大时，因为稳定性上的原因，完全不适用了. 乘幂法也是利用 Крылов 子空

间，求特征值、特征向量的办法，不过常常只能计算得个别特征值和特征向量，并且实际计算中收敛速度常很慢.

Lanczos 算法，也是利用 Крылов 子空间提供的讯息，来求 A 的近似特征值和近似特征向量的方法. 它也是将序列(15)正交化的办法. 作法是取定一个单位向量 $\boldsymbol{q}_1$，按照第 2 章 § 2 的(7)式，求得 $\boldsymbol{q}_2, \boldsymbol{q}_3, \cdots, \boldsymbol{q}_j$ 和 $\alpha_1, \alpha_2, \cdots, \alpha_j$; $\beta_1, \beta_2, \cdots, \beta_j$. 得到 $Q_j=[\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j]$,

$$\boldsymbol{T}_j=\begin{pmatrix} \alpha_1 & \beta_1 & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \ddots & \ddots & \\ & & \ddots & \ddots & \beta_{jn} \\ 0 & & & \beta_{jn} & \alpha_j \end{pmatrix}.$$

计算 T_j 的特征值，作为 A 的特征值的近似值，若 T_j 的特征向量是 $\boldsymbol{s}$，则 $Q_j\boldsymbol{s}$ 作为 A 的特征向量的近似. 具体程式如下:

1. 取 $\boldsymbol{q}_1$, $\|\boldsymbol{q}_1\|=1$, $\boldsymbol{u}_1=A\boldsymbol{q}_1$, $1\to j$;

2. $\alpha_j=\boldsymbol{q}_j^*\boldsymbol{u}_j$;

3. $\boldsymbol{r}_j=\boldsymbol{u}_j-\boldsymbol{q}_j\alpha_j$;

4. $\beta_j=\|\boldsymbol{r}_j\|$; (16)

5. 如果 β_j 充分小，则转入计算 T_j 的特征值和特征向量；否则 $\boldsymbol{q}_{j+1}=\boldsymbol{r}_j/\beta_j$;

6. $\boldsymbol{u}_{j+1}=A\boldsymbol{q}_{j+1}-\beta_j\boldsymbol{q}_j$, $j+1\to j$, 转 2;

假如不考虑舍入误差，计算是无限位小数进行的，则当计算到第 j 步，我们得到向量序列 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$、$\cdots$、$\boldsymbol{q}_j$ 和 $\boldsymbol{r}_j$，这是 $j+1$ 个正交向量组，实际上对于 $j=1$ 的情况只有 q_1 和 r_1 两个向量，有

$$(r_1, q_1) = (Aq_1 - \alpha_1 q_1, q_1) = (Aq_1, q_1) - \alpha_1(q_1, q_1),$$

按 α_1 的定义知 $(r_1, q_1) = 0$. 因此 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$、…、$\boldsymbol{q}_j$ 和 $\boldsymbol{r}_j$ 是一组正交化向量组对 $j=1$ 成立，假定对 $j=k$ 成立，即 $\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_k$ 和 $\boldsymbol{r}_k$ 是一组正交化组，假定 $\boldsymbol{r}_k \neq 0$. 考虑 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$、…、$\boldsymbol{q}_k$、$\boldsymbol{q}_{k+1}$ 和 $\boldsymbol{r}_{k+1}$. 因为 $\boldsymbol{r}_k = \beta_k \boldsymbol{q}_{k+1}$, 故 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$、…、$\boldsymbol{q}_k$、$\boldsymbol{q}_{k+1}$ 是一组单位正交向量组，为此只要证 $(\boldsymbol{r}_{k+1}, \boldsymbol{q}_l) = 0$, $l = 1, 2, \cdots, k, k+1$, 对于 $l < k$,

$$\begin{aligned}(\boldsymbol{r}_{k+1}, \boldsymbol{q}_l) &= (A\boldsymbol{q}_{k+1} - \alpha_{k+1}\boldsymbol{q}_{k+1} - \beta_k \boldsymbol{q}_k, \boldsymbol{q}_l) \\ &= (A\boldsymbol{q}_{k+1}, \boldsymbol{q}_l) = (\boldsymbol{q}_{k+1}, A\boldsymbol{q}_l) \\ &= (\boldsymbol{q}_{k+1}, \beta_l \boldsymbol{q}_{l+1} + \alpha_l \boldsymbol{q}_l + \beta_{l-1}\boldsymbol{q}_{l-1}) = 0.\end{aligned}$$

又

$$\begin{aligned}(\boldsymbol{r}_{k+1}, \boldsymbol{q}_k) &= (A\boldsymbol{q}_{k+1} - \alpha_{k+1}\boldsymbol{q}_{k+1} - \beta_k \boldsymbol{q}_k, \boldsymbol{q}_k) \\ &= (A\boldsymbol{q}_{k+1}, \boldsymbol{q}_k) - \beta_k(\boldsymbol{q}_k, \boldsymbol{q}_k) \\ &= (\boldsymbol{q}_{k+1}, A\boldsymbol{q}_k) - \beta_k \\ &= (\boldsymbol{q}_{k+1}, \beta_k \boldsymbol{q}_{k+1} - \alpha_k \boldsymbol{q}_k - \beta_{k-1}\boldsymbol{q}_{k-1}) - \beta_k \\ &= \beta_k(\boldsymbol{q}_{k+1}, \boldsymbol{q}_{k+1}) - \beta_k = 0,\end{aligned}$$

而

$$\begin{aligned}(\boldsymbol{r}_{k+1}, \boldsymbol{q}_{k+1}) &= (A\boldsymbol{q}_{k+1} - \alpha_{k+1}\boldsymbol{q}_{k+1} - \beta_k \boldsymbol{q}_k, \boldsymbol{q}_{k+1}) \\ &= (A\boldsymbol{q}_{k+1}, \boldsymbol{q}_{k+1}) - \alpha_{k+1}(\boldsymbol{q}_{k+1}, \boldsymbol{q}_{k+1}) = 0.\end{aligned}$$

于是我们证明了命题：若 β_1、β_2、…、β_{j-1} 都不为 0，则 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$、…、$\boldsymbol{q}_j$ 是单位正交向量组，$\boldsymbol{q}_1$、$\boldsymbol{q}_2$、…、$\boldsymbol{q}_j$ 与 $\boldsymbol{r}_j$ 正交.

$Q_j = [\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j]$ 有 $Q_j^* Q_j = I$, T_j 是一个对称三对角不可约矩阵，第一条超对角线上元素都为正的. 称 T_j 的特征值为 A 的 Ritz 值. 若 μ_i 是 T_j 的特征值，s_i 是对应 μ_i 的单位特征向量，称 $\boldsymbol{z}_i = Q_j \boldsymbol{s}_i$ 为 A 的 Ritz 向量.

将 $\boldsymbol{r}_j$ 表示成 $\beta_j \boldsymbol{q}_{j+1}$, 有下列定理.

定理 4.11 A, Q_j, T_j, β_j, $\boldsymbol{q}_{j+1}$ 之间成立如下的关系

$$AQ_j - Q_jT_j = \beta_j \boldsymbol{q}_{j+1}\boldsymbol{e}_j^*. \qquad (17)$$

证明 由构造过程

$$\beta_i \boldsymbol{q}_{i+1} = A\boldsymbol{q}_i - \alpha_i \boldsymbol{q}_i - \beta_{i-1}\boldsymbol{q}_{i-1},$$

得
$$A\boldsymbol{q}_i = \beta_i \boldsymbol{q}_{i+1} + \alpha_i \boldsymbol{q}_i + \beta_{i-1}\boldsymbol{q}_{i-1},$$

因此矩阵
$$AQ_j - Q_jT_j$$

除了第 j 列外，其余各列都为 0. 但它的第 j 列为

$$A\boldsymbol{q}_j - \beta_{j-1}\boldsymbol{q}_{j-1} - \alpha_j \boldsymbol{q}_j = \beta_j \boldsymbol{q}_{j+1},$$

$n\times j$ 矩阵 $\beta_j \boldsymbol{q}_{j+1}\boldsymbol{e}_j^*$ 的第 j 列为 $\beta_j \boldsymbol{q}_{j+1}$，其余各列都为 0，因此 (17) 成立. 证毕.

推论 存在 i_1、i_2、…、i_j j 个不相同的自然数，使得

$$|\lambda_{i_k} - \mu_k| \leqslant \beta_j.$$

证明 应用定理 4.8，此时 $R = \beta_j \boldsymbol{q}_{j+1}\boldsymbol{e}_j^*$，故存在 i_1、i_2、…、i_j，使得

$$|\lambda_{i_k} - \mu_k| \leqslant \|\beta_j \boldsymbol{q}_{j+1}\boldsymbol{e}_j^*\| \leqslant \beta_j. \text{ 证毕.}$$

在 (17) 两边左乘 Q_j^*，因为 Q_j 各列正交于 $\boldsymbol{q}_{j+1}$，因此有

$$Q_j^*AQ_j - Q_j^*Q_jT_j = 0,$$

$$T_j = Q_j^*AQ_j.$$

所以由定理 4.9 可知，在所有 $j\times j$ 矩阵 M_j 中，使 $R = AQ_j - Q_jM_j$ 范数 $\|R\|$ 最小的是 T_j.

定理 4.12 Ritz 向量 $\boldsymbol{z}_i = Q_j\boldsymbol{s}_i$，成立如下关系式

$$A\boldsymbol{z}_i - \mu_i \boldsymbol{z}_i = \beta_j s_{ji} \boldsymbol{q}_{j+1},$$

这里 s_{ji} 是 $\boldsymbol{s}_i$ 的第 j 个分量. 又 $\boldsymbol{z}_i$ 的 Rayleigh 商为 μ_i，即

$$\mu_i = (A\boldsymbol{z}_i, \boldsymbol{z}_i)/(\boldsymbol{z}_i, \boldsymbol{z}_i).$$

证明 在 (17) 两边右乘向量 $\boldsymbol{s}_i$，

$$AQ_j\boldsymbol{s}_i - Q_jT_j\boldsymbol{s}_i = \beta_j s_{ji}\boldsymbol{q}_{j+1},$$

$$A\boldsymbol{z}_i - \mu_i Q_j\boldsymbol{s}_i = \beta_j s_{ji}\boldsymbol{q}_{j+1},$$

$$A\boldsymbol{z}_i - \mu_i \boldsymbol{z}_i = \beta_j s_{ji}\boldsymbol{q}_{j+1},$$

又

$$(A\boldsymbol{z}_i,\ \boldsymbol{z}_i)=(AQ_j\boldsymbol{s}_i,\ Q_j\boldsymbol{s}_i)=(Q_j^*AQ_j\boldsymbol{s}_i,\ \boldsymbol{s}_i)$$
$$=(T_j\boldsymbol{s}_i,\ \boldsymbol{s}_i)=\mu_i,$$

又 $(\boldsymbol{z}_i,\ \boldsymbol{z}_i)=(Q_j\boldsymbol{s}_i,\ Q_j\boldsymbol{s}_i)=(Q_j^*Q_j\boldsymbol{s}_i,\ \boldsymbol{s}_i)=(\boldsymbol{s}_i,\ \boldsymbol{s}_i)=1,$

故 $\mu_i=(A\boldsymbol{z}_i,\ \boldsymbol{z}_i)/(\boldsymbol{z}_i,\ \boldsymbol{z}_i)=\rho(\boldsymbol{z}_i,\ A)$。证毕.

推论 若 λ 是 A 的特征值中与 μ_i 最接近的一个，若

$$d_i=\min_{\lambda_l\neq\lambda}|\lambda_l-\mu_i|, \tag{18}$$

$\boldsymbol{z}_i$ 与 $\boldsymbol{y}$ 的交角为 θ_i, $\boldsymbol{y}$ 是对应 λ 的特征向量，其意义如定理 4.4，则

$$|\lambda-\mu_i|\leqslant(\beta_j s_{ji})^2/d_i, \tag{19}$$

$$|\sin\theta_i|\leqslant|\beta_j s_{ji}|/d_i. \tag{20}$$

(19)，(20)两式给出了 A 的Ritz 值和 Ritz 向量，对于 A 的特征值和特征向量逼近程度的估计. 逼近依赖于量 $|\beta_j s_{ji}|$，通常将它记为 β_{ji}，如果 β_{ji} 很小，那么逼近就会好，而从分母上的 d_i，可以知道，如果 λ 是 A 的特征值中孤立的一个，d_i 就会大，逼近就会好. 反之如果 λ 是 A 的特征值的密集丛中的一个 d_i 就会小，于是 β_{ji}/d_i 就会大，估计式的右端就大，实际计算经验也告诉人们，此时实际逼近也是比较差的.

现在来分析一下 β_j，从

$$\boldsymbol{r}_i=A\boldsymbol{q}_i-\alpha_i\boldsymbol{q}_i-\beta_{i-1}\boldsymbol{q}_{i-1},$$

可以知道 $\boldsymbol{r}_i$ 是向量 $A\boldsymbol{q}_i$，在子空间 $\{\boldsymbol{q}_1,\ \boldsymbol{q}_2,\ \cdots,\ \boldsymbol{q}_i\}$ 中投影(最佳逼近)的偏差，因为$\{\boldsymbol{q}_1,\ \boldsymbol{q}_2,\ \cdots,\ \boldsymbol{q}_i\}$就是 $\{\boldsymbol{q}_1,\ A\boldsymbol{q}_1,\ \cdots,\ A^{i-1}\boldsymbol{q}_1\}$，因此如果 $\boldsymbol{q}_1$ 对于矩阵 A 的最小化零多项式是m次多项式，则 $\boldsymbol{q}_1$、$A\boldsymbol{q}_1$、…、$A^{m-1}\boldsymbol{q}_1$ 是线性无关的，而 $\boldsymbol{q}_1$、$A\boldsymbol{q}_1$、…、$A^{m-1}\boldsymbol{q}_1$、$A^m\boldsymbol{q}_1$ 是线性相关的，于是对于这样的 q_1，

$$\boldsymbol{r}_l=A\boldsymbol{q}_l-\alpha_l\boldsymbol{q}_l-\beta_{l-1}\boldsymbol{q}_{l-1}\neq0,\quad l<m,$$
$$\boldsymbol{r}_m=A\boldsymbol{q}_m-\alpha_m\boldsymbol{q}_m-\beta_{m-1}\boldsymbol{q}_{m-1}=0,$$

即 $\beta_l>0$ 当 $1\leqslant l<m$，而 $\beta_m=0$. 此时 $\{q_1, q_2, \cdots, q_m\}$ 是不变子空间，对应的 Ritz 值和 Ritz 向量，恰为 A 的特征值和特征向量.

当 n 很大时，m 可能也很大，对于比较大的 j, $j<m$, β_j 是否会很小？从实际计算的经验来看，除了少数情况外没有迹象显示 β_j 会十分小. 而 $|s_{ji}|$ 当 j 足够大时却常常是十分小的，尤其是对于 T_j 的两端特征值，所对应的特征向量的 $|s_{ji}|$，因此 β_{ji} 常常可以达到很小，但这一点理论上还不能给出证明.

下面来讨论 T_j 的特征值与 T_k 的特征值之间的关系，$j<k$. 有

定理 4.13 设 $\mu_i^{(j)}$ 是 T_j 的第 i 个特征值，当 $j<k$ 时，可以找到 T_k 的一个特征值 $\mu^{(k)}$，使得

$$|\mu_i^{(j)}-\mu^{(k)}|\leqslant\beta_{ji}.$$

证明 设 s_i 是 T_j 对应 $\mu_i^{(j)}$ 的单位特征向量. 造一个 k 维列向量 $s_i'=\begin{pmatrix} s_i \\ 0 \end{pmatrix}$，也有 $\|s_i'\|=1$，我们来估计

$$\|T_k s_i'-\mu_i^{(j)}s_i'\|,$$

$$T_k s_i'-\mu_i^{(j)}s_i'=\begin{pmatrix} T_j & \begin{matrix} 0 \\ \beta_j \end{matrix} \\ \begin{matrix} 0 & \beta_j \end{matrix} & T_{j+1,k} \end{pmatrix}\begin{pmatrix} s_i \\ 0 \end{pmatrix}-\mu_i^{(j)}\begin{pmatrix} s_i \\ 0 \end{pmatrix}$$

$$=\begin{pmatrix} T_j s_i-\mu_i^{(j)}s_i \\ \beta_j s_{ji} \\ 0 \end{pmatrix}=\begin{pmatrix} 0 \\ \beta_j s_{ji} \\ 0 \end{pmatrix},$$

右端是一个只有第 $j+1$ 个分量为 $\beta_j s_{ji}$，而其余分量都为 0 的向量. 因此

$$\|T_k s_i'-\mu_i^{(j)}s_i'\|=\beta_{ji},$$

利用定理 4.3，就知存在 T_k 的特征值 $\mu^{(k)}$ 使得

$$|\mu_i^{(j)}-\mu^{(k)}|\leqslant\beta_{ji}. \quad 证毕.$$

这个定理告诉我们，如果计算到某步 j，T_j 的某个特征值 $\mu_i^{(j)}$，它的 β_{ji} 很小，那么 $\mu_i^{(j)}$ 是 A 的特征值 λ 很好的近似值，往后 T_k 的特征值中，一定也有一个很好地逼近 λ. 同时也告诉我们，如果 T_j 的特征值 $\mu^{(j)}$ 与 T_k 的所有特征值都差别很大，那么 $\mu^{(j)}$ 不会是 A 的特征值的近似值. 因此这个定理，对实际计算有指导意义.

我们以 $L^j(A,\ \boldsymbol{q}_1)=(Q_j,\ T_j)$ 表示对于对称矩阵 A，单位向量 $\boldsymbol{q}_1$，通过 Lanczos 算法计算到 j 步，产生出矩阵 $Q_j=[\boldsymbol{q}_1,\ \boldsymbol{q}_2,\ \cdots,\ \boldsymbol{q}_j]$ 和 T_j.

定理 4.14 如果 $L^j(A,\ \boldsymbol{q}_1)=(Q_j,\ T_j)$，则

1. $L^j(\alpha A,\ \boldsymbol{q}_1)=(Q_j,\ \alpha T_j)$ 对所有正数 α;
2. $L^j(A-\alpha I,\ \boldsymbol{q}_1)=(Q_j,\ T_j-\alpha I)$ 对所有实数 α;
3. $L^j(P^*AP,\ P^*\boldsymbol{q}_1)=(P^*Q_j,\ T_j)$ 对所有 $n\times n$ 正交阵 P.

证明 1. 记 $B=\alpha A$，由 B，$\boldsymbol{q}_1$ 通过 Lanczos 算法，得到 $\tilde{\boldsymbol{q}}_2$、$\tilde{\boldsymbol{q}}_3$、$\cdots$、$\tilde{\boldsymbol{q}}_j$、$\tilde{\alpha}_1$、$\tilde{\beta}_1$、$\cdots$、$\tilde{\alpha}_{j-1}$、$\tilde{\beta}_{j-1}$、$\tilde{\alpha}_j$，于是有

$$\tilde{\beta}_1\tilde{\boldsymbol{q}}_2=B\boldsymbol{q}_1-\tilde{\alpha}_1\boldsymbol{q}_1,$$

其中
$$\tilde{\alpha}_1=(B\boldsymbol{q}_1,\ \boldsymbol{q}_1)=\alpha(A\boldsymbol{q}_1,\ \boldsymbol{q}_1)=\alpha\alpha_1.$$

$$\tilde{\beta}_1=\|B\boldsymbol{q}_1-\tilde{\alpha}_1\boldsymbol{q}_1\|=\alpha\|A\boldsymbol{q}_1-\alpha_1\boldsymbol{q}_1\|=\alpha\beta_1,$$

于是
$$\tilde{\beta}_1\tilde{\boldsymbol{q}}_2=B\boldsymbol{q}_1-\tilde{\alpha}_1\boldsymbol{q}_1,$$

即为
$$\alpha\beta_1\tilde{\boldsymbol{q}}_2=\alpha(A\boldsymbol{q}_1-\alpha_1\boldsymbol{q}_1)=\alpha\beta_1\boldsymbol{q}_2,$$

由此可知 $\tilde{\boldsymbol{q}}_2=\boldsymbol{q}_2$. 同理可以证明 $\tilde{\boldsymbol{q}}_3$, $\cdots$, $\tilde{\boldsymbol{q}}_j$ 依次为 $\boldsymbol{q}_3$、$\cdots$、$\boldsymbol{q}_j$; 而 $\tilde{\alpha}_2$、$\tilde{\beta}_2$、$\cdots$、$\tilde{\alpha}_{j-1}$、$\tilde{\beta}_{j-1}$、$\tilde{\alpha}_j$ 依次为 $\alpha\alpha_2$、$\alpha\beta_2$、$\cdots$、$\alpha\alpha_{j-1}$、$\alpha\beta_{j-1}$、$\alpha\alpha_j$ 即得 $L^j(\alpha A,\ \boldsymbol{q}_1)=(Q_j,\ \alpha T_j)$.

2 和 3 的证明类同，不再重复. 证毕.

§3 Kaniel-Paige-Saad 理论

在上节分析中知道：T_j 的特征值可以作为 A 的某个特征值的近似值，但不知道是 A 的第几个特征值的近似值. S. Kaniel 在 1966 年发表的文章[25]中，给出了 $\mu_i-\lambda_i$ 的估计，不过证明中有些错误，C. Paige 在 1971 年的学位论文[13]中，重新研究了这个问题，给出了 $\mu_i-\lambda_i$ 的估计，并且也给出了 $\sin\phi_i$ 的估计，这里 ϕ_i 表示 Ritz 向量 $\boldsymbol{z}_i$ 与 A 的特征向量 $\boldsymbol{y}_i$ 的交角. Y. Saad 在 1980 年发表的文章[26]中，对于 $\mu_i-\lambda_i$ 给出比 Paige 的更简单的估计. 所以这些结果，我们统称为 Kaniel-Paige-Saad 理论.

本节大部分定理的证明思想取自 B. N. Parlett 的书[10].

首先来进一步研究 A 的从 T_j 得到的 Ritz 向量，$\boldsymbol{z}_1$、$\boldsymbol{z}_2$、…、$\boldsymbol{z}_j$ 的性质.

1. 正交性

$$(\boldsymbol{z}_i,\ \boldsymbol{z}_l)=\delta_{il}. \tag{21}$$

因为 $\boldsymbol{z}_i=Q_j\boldsymbol{s}_i$，$\boldsymbol{z}_l=Q_j\boldsymbol{s}_l$，故

$$(\boldsymbol{z}_i,\ \boldsymbol{z}_l)=(Q_j\boldsymbol{s}_i,\ Q_j\boldsymbol{s}_l)=(Q_j^*Q_j\boldsymbol{s}_i,\ \boldsymbol{s}_l)=(\boldsymbol{s}_i,\ \boldsymbol{s}_l)=\delta_{il}.$$

2. $$\{\boldsymbol{q}_1,\ \boldsymbol{q}_2,\ \cdots,\ \boldsymbol{q}_j\}=\{\boldsymbol{z}_1,\ \boldsymbol{z}_2,\ \cdots,\ \boldsymbol{z}_j\}. \tag{22}$$

记 $[\boldsymbol{s}_1,\ \boldsymbol{s}_2,\ \cdots,\ \boldsymbol{s}_j]=S$，有 $\det(S)\neq 0$，从

$$Q_jS=[\boldsymbol{z}_1,\ \boldsymbol{z}_2,\ \cdots,\ \boldsymbol{z}_j],$$

可知(22)成立.

3. A-正交性

$$(A\boldsymbol{z}_l,\ \boldsymbol{z}_k)=\begin{cases}\mu_l, & l=k,\\ 0, & l\neq k.\end{cases}$$

这是因为

$$(A\boldsymbol{z}_l, \boldsymbol{z}_k) = (Q_j^* A Q_j \boldsymbol{s}_l, \boldsymbol{s}_k)$$
$$= (T_j \boldsymbol{s}_l, \boldsymbol{s}_k) = \mu_l (\boldsymbol{s}_l, \boldsymbol{s}_k) = \mu_l \delta_{lk}.$$

4．记 P_j 表示 j 次多项式全体，$m \in P_j$，则

$$(m(A)\boldsymbol{q}_1, \boldsymbol{z}_l) = 0$$

的充要条件是 $(\lambda - \mu_l) | m(\lambda)$.

证明 先证充分性．假设 $(\lambda - \mu_l) | m(\lambda)$，有

$$m(\lambda) = (\lambda - \mu_l) g(\lambda), \quad g(\lambda) \in P_{j-1},$$
$$(m(A)\boldsymbol{q}_1, \boldsymbol{z}_l) = (g(A)\boldsymbol{q}_1, (A - \mu_l)\boldsymbol{z}_l),$$
$$g(A)\boldsymbol{q}_1 \in \{\boldsymbol{q}_1, A\boldsymbol{q}_1, \cdots, A^{j-1}\boldsymbol{q}_1\} = \{\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j\},$$

故
$$g(A)\boldsymbol{q}_1 = Q_j \boldsymbol{h},$$

h 是一个 j 维向量．故

$$(m(A)\boldsymbol{q}_1, \boldsymbol{z}_l) = (Q_j \boldsymbol{h}, (A - \mu_l) Q_j \boldsymbol{s}_l)$$
$$= (\boldsymbol{h}, (T_j - \mu_l)\boldsymbol{s}_l) = 0.$$

再证必要性．若 $(m(A)\boldsymbol{q}_1, \boldsymbol{z}_l) = 0$,

$$m(\lambda) = (\lambda - \mu_l) g(\lambda) + d,$$

于是

$$(m(A)\boldsymbol{q}_1, \boldsymbol{z}_l) = (((A - \mu_l I) g(A) + dI)\boldsymbol{q}_1, \boldsymbol{z}_l)$$
$$= ((A - \mu_l) g(A)\boldsymbol{q}_1, \boldsymbol{z}_l) + d(\boldsymbol{q}_1, \boldsymbol{z}_l)$$
$$= d(\boldsymbol{q}_1, Q_j \boldsymbol{s}_l) = d(\boldsymbol{e}_1, \boldsymbol{s}_l) = d s_{1l},$$

但由第 2 章的定理 2.9 知 S_l 的第一个分量 $s_{1l} \neq 0$，故必须 $d = 0$.

类似于 Ritz 向量的性质 4，对于特征向量 $\boldsymbol{y}_l$ 也有如下性质.

5．若 $(\boldsymbol{q}_1, \boldsymbol{y}_l) \neq 0$，$m(\lambda) \in P_j$，则

$$(m(A)\boldsymbol{q}_1, \boldsymbol{y}_l) = 0$$

的充分必要条件是

$$\lambda - \lambda_l | m(\lambda).$$

证明 充分性. 若$\lambda-\lambda_l|m(\lambda)$, 即$m(\lambda)=(\lambda-\lambda_l)g(\lambda)$, $g(\lambda)\in P_{j-1}$, 由此

$$(m(A)\boldsymbol{q}_1, \boldsymbol{y}_l)=((A-\lambda_l I)g(A)\boldsymbol{q}_1, \boldsymbol{y}_l)$$
$$=(g(A)\boldsymbol{q}_1, (A-\lambda_l I)\boldsymbol{y}_l)=0.$$

必要性. 若$(m(A)\boldsymbol{q}_1, \boldsymbol{y}_l)=0$, 如果$m(\lambda)=(\lambda-\lambda_l)g(\lambda)+d$, 则

$$((A-\lambda_l I)g(A)\boldsymbol{q}_1+d\boldsymbol{q}_1, \boldsymbol{y}_l)$$
$$=((A-\lambda_l I)g(A)\boldsymbol{q}_1, \boldsymbol{y}_l)+d(\boldsymbol{q}_1, \boldsymbol{y}_l)=d(\boldsymbol{q}_1, \boldsymbol{y}_l)=0,$$

但$(\boldsymbol{q}_1, \boldsymbol{y}_l)\neq 0$, 故必须$d=0$, 即$(\lambda-\lambda_l)|m(\lambda)$.

定理4.15 设μ_l是T_j的从小到大排列的第l个特征值, λ_l是A的从小到大排列的第l个特征值, 则

$$\mu_l-\lambda_l\geqslant 0,\quad \lambda_{n-l+1}-\mu_{j-l+1}\geqslant 0,\ l=1, 2, \cdots, j.$$

证明 将$Q_j=[\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_j]$扩充成一个正交矩阵$Q=(Q_j, \widetilde{Q}_{n-j})$,

$$Q^*AQ=\begin{pmatrix} Q_j^*AQ_j & Q_j^*A\widetilde{Q}_{n-j} \\ \widetilde{Q}_{n-j}^*AQ_j & \widetilde{Q}_{n-j}^*A\widetilde{Q}_{n-j}\end{pmatrix}$$
$$=\begin{pmatrix} T_j & Q_j^*A\widetilde{Q}_{n-j} \\ \widetilde{Q}_{n-j}^*AQ_j & \widetilde{Q}_{n-j}^*A\widetilde{Q}_{n-j}\end{pmatrix},$$

记$B=Q^*AQ$, 它的特征值与A的相同, 利用极大极小原理知

$$\lambda_l=\min_{V_l}\max_{x\in V_l}(B\boldsymbol{x}, \boldsymbol{x})/(\boldsymbol{x}, \boldsymbol{x}),$$

特别取一个l维子空间$\widetilde{V}_l$, 它是由最后$n-j$个分量为0的向量构成的, 因为$l\leqslant j$, 因此这样的子空间是存在的, 当$l<j$时, 这样的子空间可以有很多, 但它们的全体总是一般l维子空间的全体构成的集合中的子集. 因此

$$\lambda_l\leqslant\min_{\widetilde{V}_l}\max_{x\in\widetilde{V}_l}(B\boldsymbol{x}, \boldsymbol{x})/(\boldsymbol{x}, \boldsymbol{x})$$
$$=\min_{\widetilde{V}_l}\max_{y\in\widetilde{V}_l}(T_j\boldsymbol{y}, \boldsymbol{y})/(\boldsymbol{y}, \boldsymbol{y})=\mu_l.$$

同样可证

$$\lambda_{n-l+1}-\mu_{j-l+1}\geqslant 0.$$ 证毕.

下面介绍 Kaniel–Paige 理论.

定理 4.16 设 $\phi_i=\angle(\boldsymbol{y}_i,\ \boldsymbol{z}_i)$, $i=1, 2, \cdots, j$, 则

$$\sin^2\phi_i\leqslant\left[\mu_i-\lambda_i+\sum_{k=1}^{i-1}(\lambda_{i+1}-\lambda_k)\sin^2\phi_k\right]\Big/(\lambda_{i+1}-\lambda_i). \tag{23}$$

证明

$$\boldsymbol{z}_i=\sum_{k=1}^{n}r_{ik}\boldsymbol{y}_k,\quad \sum_{k=1}^{n}r_{ik}^2=1,$$

$$(\boldsymbol{y}_i,\ \boldsymbol{z}_i)=r_{ii}=\cos\phi_i.$$

$$\sin^2\phi_i=\sum_{\substack{k=1\\k\neq i}}^{n}r_{ik}^2,$$

$$\mu_i=\boldsymbol{z}_i^*A\boldsymbol{z}_i=\sum_{k=1}^{n}r_{ik}^2\lambda_k,$$

$$\lambda_i=\sum_{k=1}^{n}r_{ik}^2\lambda_i,$$

所以

$$\mu_i-\lambda_i=\sum_{\substack{k=1\\k\neq i}}^{n}r_{ik}^2(\lambda_k-\lambda_i),$$

或

$$\mu_i-\lambda_i+\sum_{k=1}^{i-1}r_{ik}^2(\lambda_i-\lambda_k)=\sum_{k=i+1}^{n}r_{ik}^2(\lambda_k-\lambda_i)$$

$$\geqslant(\lambda_{i+1}-\lambda_i)\sum_{k=i+1}^{n}r_{ik}^2,$$

不等式两边都加上项$(\lambda_{i+1}-\lambda_i)\sum\limits_{k=1}^{i-1}r_{ik}^2$, 即得

$$\sin^2\phi_i\leqslant\left[\mu_i-\lambda_i+\sum_{k=1}^{i-1}r_{ik}^2(\lambda_{i+1}-\lambda_k)\right]\Big/(\lambda_{i+1}-\lambda_i), \tag{24}$$

另一方面, 将特征向量 y_k 在 Ritz 向量 z_k 和跟 z_k 的正交方向上作分解,

$$\boldsymbol{y}_k=\boldsymbol{z}_k\cos\phi_k+\boldsymbol{h}_k\sin\phi_k,\quad k=1,\ 2,\ \cdots,\ j,$$

$\boldsymbol{h}_k$ 是与 $\boldsymbol{z}_k$ 正交方向上的单位向量，当 $i\neq k$ 时有

$$r_{ik}=(\boldsymbol{z}_i,\ \boldsymbol{y}_k)=(\boldsymbol{z}_i,\ \boldsymbol{y}_k-\boldsymbol{z}_k\cos\phi_k)=(\boldsymbol{z}_i,\ \boldsymbol{h}_k\sin\phi_k),$$

从而知道
$$r_{ik}^2\leqslant\sin^2\phi_k,$$

将此代入(24)式，即得

$$\sin^2\phi_i\leqslant\left[\ \mu_i-\lambda_i+\sum_{k=1}^{i-1}(\lambda_{i+1}-\lambda_k)\sin^2\phi_k\ \right]\Big/(\lambda_{i+1}-\lambda_i).$$

证毕.

定理 4.17 对任意自然数 $l\leqslant j$，任意非零向量 $\boldsymbol{s}\in\{\boldsymbol{q}_1,\ A\boldsymbol{q}_1,\ \cdots,\ A^{j-1}\boldsymbol{q}_1\}$，如果满足

$$\boldsymbol{s}^*\boldsymbol{y}_k=0,\quad k=1,\ 2,\ \cdots,\ l-1,$$

则
$$\mu_l\leqslant\rho(\boldsymbol{s})+\sum_{k=1}^{l-1}(\lambda_n-\mu_k)\sin^2\phi_k$$
$$\leqslant\rho(\boldsymbol{s})+\sum_{k=1}^{l-1}(\lambda_n-\lambda_k)\sin^2\phi_k,$$

其中
$$\rho(\boldsymbol{s})=(A\boldsymbol{s},\ \boldsymbol{s})/(\boldsymbol{s},\ \boldsymbol{s}).$$

证明 结论中第二个不等式成立是明显的，只要利用定理 4.15 即得.

将 $\boldsymbol{s}$ 分解成

$$\boldsymbol{s}=\boldsymbol{t}+\sum_{k=1}^{l-1}r_k\boldsymbol{z}_k,$$

其中 $\boldsymbol{t}$ 是与 $\boldsymbol{z}_k(k=1,\ 2,\ \cdots,\ l-1)$ 正交的向量.

$$(\boldsymbol{s},\ \boldsymbol{s})=(\boldsymbol{t},\ \boldsymbol{t})+\sum_{k=1}^{l-1}r_k^2,$$

$$r_k=(\boldsymbol{s},\ \boldsymbol{z}_k)=(\boldsymbol{s},\ \boldsymbol{z}_k-\boldsymbol{y}_k\cos\phi_k),\quad k=1,\ 2,\ \cdots,\ l-1,$$

所以
$$r_k^2\leqslant\|\boldsymbol{s}\|^2\sin^2\phi_k,$$

$$(A\boldsymbol{s},\ \boldsymbol{s})=(A\boldsymbol{t},\ \boldsymbol{t})+\left(\sum_{k=1}^{l-1}r_kA\boldsymbol{z}_k,\ \boldsymbol{t}\right)$$
$$+\left(A\boldsymbol{t},\ \sum_{k=1}^{l-1}r_k\boldsymbol{z}_k\right)+\left(\sum_{k=1}^{l-1}r_kA\boldsymbol{z}_k,\ \sum_{s=1}^{l-1}r_s\boldsymbol{z}_s\right),$$

因为 $\boldsymbol{t}\in\{\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_j\}$，故利用 $\boldsymbol{z}_l$ 的性质 3 知

$$(A\boldsymbol{t}, \boldsymbol{z}_k)=(A\boldsymbol{z}_k, \boldsymbol{t})=0, \quad k=1, 2, \cdots, l-1,$$

因此

$$\rho(s)=\left[(A\boldsymbol{t}, \boldsymbol{t})+\sum_{k=1}^{l-1}r_k^2(A\boldsymbol{z}_k, \boldsymbol{z}_k)\right]\Big/(\boldsymbol{s}, \boldsymbol{s})$$
$$=\left[(A\boldsymbol{t}, \boldsymbol{t})+\sum_{k=1}^{l-1}\mu_k r_k^2\right]\Big/(\boldsymbol{s}, \boldsymbol{s}),$$

两边减去 λ_n，利用 $(\boldsymbol{t}, \boldsymbol{t})+\sum_{k=1}^{l-1}r_k^2=(\boldsymbol{s}, \boldsymbol{s})$，得

$$\rho(\boldsymbol{s})-\lambda_n=\left[((A-\lambda_n I)\boldsymbol{t}, \boldsymbol{t})+\sum_{k=1}^{l-1}(\mu_k-\lambda_n)r_k^2\right]\Big/(\boldsymbol{s},\boldsymbol{s}),$$
$$((A-\lambda_n I)\boldsymbol{t}, \boldsymbol{t})\leqslant 0, \quad (\boldsymbol{s}, \boldsymbol{s})\geqslant(\boldsymbol{t}, \boldsymbol{t}),$$

故

$$\rho(\boldsymbol{s})-\lambda_n\geqslant\frac{((A-\lambda_n I)\boldsymbol{t}, \boldsymbol{t})}{(\boldsymbol{t}, \boldsymbol{t})}+\left(\sum_{k=1}^{l-1}(\mu_k-\lambda_n)r_k^2\right)\Big/(\boldsymbol{s}, \boldsymbol{s}),$$

因为 $\boldsymbol{t}$ 正交于 $\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_{l-1}$，故

$$\frac{((A-\lambda_n I)\boldsymbol{t}, \boldsymbol{t})}{(\boldsymbol{t}, \boldsymbol{t})}\geqslant\mu_l-\lambda_n,$$

由此得到

$$\mu_l-\lambda_n\leqslant\rho(\boldsymbol{s})-\lambda_n+\left(\sum_{k=1}^{l-1}(\lambda_n-\mu_k)r_k^2\right)\Big/(\boldsymbol{s}, \boldsymbol{s}),$$

再将 $r_k^2\leqslant\|s\|^2\sin^2\phi_k$ 代入，即得

$$\mu_l\leqslant\rho(\boldsymbol{s})+\sum_{k=1}^{l-1}(\lambda_n-\mu_k)\sin^2\phi_k. \quad \text{证毕.}$$

推论 设非零向量 $s\in\{\boldsymbol{q}_1, A\boldsymbol{q}_1, \cdots, A^{j-1}\boldsymbol{q}_1\}$，$(\boldsymbol{s}, \boldsymbol{y}_k)=0$，$k=1, 2, \cdots, l-1$，

$$\rho(\boldsymbol{s}, A-\lambda_l I)=((A-\lambda_l I)\boldsymbol{s}, \boldsymbol{s})/(\boldsymbol{s}, \boldsymbol{s}),$$

则

$$\mu_l-\lambda_l\leqslant\rho(\boldsymbol{s}, A-\lambda_l I)+\sum_{k=1}^{l-1}(\lambda_n-\lambda_k)\sin^2\phi_k. \qquad (25)$$

证明 只要在定理中用 $A-\lambda_l I$ 代替 A，即得.

从定理 4.16 的(23)可知如果 $\mu_k-\lambda_k(k=1, 2, \cdots, l-1)$ 是很小的量，那么 $\sin^2\phi_k(k=1, 2, \cdots, l-1)$ 也是很小的量，当然这里假定 $\lambda_{k+1}-\lambda_k$ 都不是十分小，在此再利用(25)式，如果 $\rho(\boldsymbol{s}, A-\lambda_l I)$ 也是很小量的话，那么 $\mu_l-\lambda_l$ 也是小量. 因此要对 $\rho(\boldsymbol{s}, A-\lambda_l I)$ 来进行分析.

定理 4.18 设 $\boldsymbol{q}_1=\sum_{i=1}^{n}\alpha_i\boldsymbol{y}_i$，对 $l\leqslant j$，$\alpha_l\neq 0$，记

$$\boldsymbol{h}=\sum_{i=l+1}^{n}\alpha_i\boldsymbol{y}_i\Big/\left(\sum_{i=l+1}^{n}\alpha_i^2\right)^{\frac{1}{2}},$$

则对任意 $m(\lambda)\in P_{j-1}$，有

$$\rho(m(A)\boldsymbol{q}_1, A-\lambda_l I)$$

$$\leqslant(\lambda_n-\lambda_l)\left[\frac{\left(\sum_{i=l+1}^{n}\alpha_i^2\right)^{\frac{1}{2}}}{\alpha_l}\frac{\|m(A)\boldsymbol{h}\|}{m(\lambda_l)}\right]^2. \tag{26}$$

证明

$$\boldsymbol{q}_1=\sum_{i=1}^{l}\alpha_i\boldsymbol{y}_i+\boldsymbol{h}\left(\sum_{i=l+1}^{n}\alpha_i^2\right)^{\frac{1}{2}},$$

$$m(A)\boldsymbol{q}_1=\sum_{i=1}^{l}\alpha_i m(\lambda_i)\boldsymbol{y}_i+m(A)\boldsymbol{h}\left(\sum_{i=l+1}^{n}\alpha_i^2\right)^{\frac{1}{2}},$$

$$((A-\lambda_l I)m(A)\boldsymbol{q}_1, m(A)\boldsymbol{q}_1)$$

$$=\sum_{i=1}^{l}\alpha_i^2 m(\lambda_i)^2(\lambda_i-\lambda_l)$$

$$+((A-\lambda_l I)m(A)\boldsymbol{h}, m(A)\boldsymbol{h})\left(\sum_{i=l+1}^{n}\alpha_i^2\right),$$

因为 $\lambda_i-\lambda_l\leqslant 0$，因此

$$\rho(m(A)\boldsymbol{q}_1, (A-\lambda_l I))\leqslant((A-\lambda_l I)m(A)\boldsymbol{h}, m(A)\boldsymbol{h})$$

$$\times\left(\sum_{i=l+1}^{n}\alpha_i^2\right)\Big/(m(A)\boldsymbol{q}_1, m(A)\boldsymbol{q}_1),$$

但

$$(m(A)\boldsymbol{q}_1,\ m(A)\boldsymbol{q}_1)\geqslant(\alpha_l m(\lambda_l))^2,$$

$$((A-\lambda_l I)m(A)\boldsymbol{h},\ m(A)\boldsymbol{h})\leqslant(\lambda_n-\lambda_l)\|m(A)\boldsymbol{h}\|^2,$$

合并这些结果，即得(26)．证毕．

下面给出 $\mu_l-\lambda_l$ 的估计．

定理 4.19 对于 $l\leqslant j$

$$\mu_l-\lambda_l\leqslant(\lambda_n-\lambda_l)\frac{\sum\limits_{k=l+1}^{n}\alpha_k^2}{\alpha_l^2}\left(\frac{\prod\limits_{k=1}^{l-1}\frac{\lambda_k-\lambda_n}{\lambda_k-\lambda_l}}{T_{j-l}(1+2r)}\right)^2+\sum_{k=1}^{l-1}(\lambda_n-\lambda_k)\sin^2\phi_k, \tag{27}$$

这里 $T_k(x)$ 是 k 阶 Чебышев 多项式，$r=\dfrac{\lambda_l-\lambda_{l+1}}{\lambda_{l+1}-\lambda_n}$.

证明 定理 4.17 的推论告诉我们，只要向量 $\boldsymbol{s}$，满足 1. $\boldsymbol{s}\in\{\boldsymbol{q}_1,\ A\boldsymbol{q}_1,\ \cdots,\ A^{j-1}\boldsymbol{q}_1\}$，2. $\boldsymbol{s}^*\boldsymbol{y}_k=0$，$k=1,2,\cdots,l-1$，就有

$$\mu_l-\lambda_l\leqslant\rho(\boldsymbol{s},\ A-\lambda_l I)+\sum_{k=1}^{l-1}(\lambda_n-\lambda_k)\sin^2\phi_k,$$

使用定理 4.18，取 $\boldsymbol{s}=m(A)\boldsymbol{q}_1$，自然满足条件 1，为了使它满足条件 2，利用 Ritz 向量的性质 5，知必须取

$$m(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_{l-1})g(\lambda),$$

其中 $g(\lambda)$ 是 $j-l$ 次多项式．对于这样的多项式，定理 4.17 中的

$$\|m(A)\boldsymbol{h}\|\leqslant(\lambda_n-\lambda_1)(\lambda_n-\lambda_2)\cdots(\lambda_n-\lambda_{l-1})\|g(A)\boldsymbol{h}\|,$$

$$m(\lambda_l)=(\lambda_l-\lambda_1)(\lambda_l-\lambda_2)\cdots(\lambda_l-\lambda_{l-1})g(\lambda_l),$$

$\boldsymbol{h}\in\{\boldsymbol{y}_{l+1},\ \boldsymbol{y}_{l+2},\ \cdots,\ \boldsymbol{y}_n\}$，因此若 $\boldsymbol{h}=\sum\limits_{i=l+1}^{n}\beta_i\boldsymbol{y}_i$，有

$$g(A)\boldsymbol{h}=\sum_{i=l+1}^{n}\beta_i g(\lambda_i)\boldsymbol{y}_i,$$

$$\|g(A)\boldsymbol{h}\|^2=\sum_{i=l+1}^{n}\beta_i^2(g(\lambda_i))^2,$$

我们取一个多项式 $g(\lambda)$，使得 $g(\lambda_i)^2(i=l+1,\cdots,n)$ 是小的，而 $g(\lambda_l)$ 是很大的，那么就能得到好的估计．取

$$g(\lambda)=T_{j-l}\left(\frac{2\lambda-\lambda_{l+1}-\lambda_n}{\lambda_n-\lambda_{l+1}}\right)$$

就是这样的一个，此时

$$g(\lambda_i)^2\leqslant 1\quad(i=l+1,\ l+2,\ \cdots,\ n),$$

故 $\|g(A)\boldsymbol{h}\|\leqslant 1$，而 $g(\lambda_l)^2=T_{j-l}(1+2r)^2$，$r=\dfrac{\lambda_{l+1}-\lambda_l}{\lambda_n-\lambda_{l+1}}$ 随 r 增大而增大很快．对于这样 $g(\lambda)$，有

$$\begin{aligned}\mu_l-\lambda_l\leqslant&(\lambda_n-\lambda_l)\left(\sum_{k=l+1}^{n}\alpha_k^2/\alpha_l^2\right)\\&\times\left(\prod_{k=1}^{l-1}\frac{\lambda_k-\lambda_n}{\lambda_k-\lambda_l}\Big/T_{j-l}(1+2r)\right)^2\\&+\sum_{k=1}^{l-1}(\lambda_n-\lambda_k)\sin^2\phi_k.\end{aligned}$$ 证毕．

利用(23)式和(27)式，有时可以得到 $\mu_l-\lambda_l$ 的估计．例如：已知 $\lambda_1=0$，$\lambda_2=0.01$，$\lambda_3=0.04$，λ_4，λ_5，$\cdots$，λ_n 都在 $[0.1,1]$ 中，而

$$\alpha_l=\boldsymbol{q}_1^*\cdot\boldsymbol{y}_l=0.01,\quad l=1,\ 2,\ 3,$$

取 $j=53$．

对于 $l=1$，

$$\sum_{k=l+1}^{n}\alpha_k^2=1-(\alpha_1)^2=1-0.0001=0.9999,$$

$$r=\frac{\lambda_2-\lambda_1}{\lambda_n-\lambda_2}\leqslant\frac{0.01}{0.99}=\frac{1}{99},$$

$$T_{j-l}(1+2r)=T_{52}\left(1+\frac{2}{99}\right)=1.73\times10^4,$$

$$\mu_1-\lambda_1\leqslant\frac{0.9999}{0.0001}\left[\frac{1}{T_{52}(1+2r)}\right]^2\leqslant 3.34\times10^{-5},$$

$$\sin^2\phi_1\leqslant\frac{\mu_1-\lambda_1}{\lambda_2-\lambda_1}\leqslant 3.34\times10^{-3}.$$

对于 $l=2$,

$$\sum_{k=l+1}^{n}\alpha_k^2=1-\alpha_1^2-\alpha_2^2=0.9998,$$

$$r=\frac{\lambda_3-\lambda_2}{\lambda_n-\lambda_3}\leqslant\frac{0.03}{0.96}=0.03125,$$

$$T_{51}(1+2\times0.03125)=3.39\times10^7,$$

$$\mu_2-\lambda_2\leqslant\frac{0.9998}{0.0001}\times\left[\frac{1}{0.01}\Big/T_{51}(1+0.0625)\right]^2$$
$$+\sin^2\phi_1\leqslant 3.35\times10^{-3},$$

$$\begin{aligned}\sin^2\phi_2&\leqslant\frac{\mu_2-\lambda_2}{\lambda_3-\lambda_2}+\sin^2\phi_1\frac{\lambda_3-\lambda_1}{\lambda_3-\lambda_2}\\&=\frac{3.35\times10^{-3}}{0.03}+3.34\times10^{-3}\frac{0.04}{0.03}\\&=1.16\times10^{-1}.\end{aligned}$$

依此类推，可以获得 $\mu_l-\lambda_l$ 和 $\sin^2\phi_l$ 的估计. 当然这种估计,离开实用还是很远,因为 α_l 不易知道，λ_l 也是不知道的量. 但是从理论上可以看到,当 $j\to\infty$ 时,$\mu_l-\lambda_l$ 和 $\sin^2\phi_l$ 都会趋于 0, 因此是有理论价值的. 同时也告诉我们: Lanczos 方法，首先求出来的是端部的特征值，因为 $\mu_l-\lambda_l$ 和 $\mu_{l+1}-\lambda_{l+1}$, 分别依赖于 $T_{j-l}(1+2r)$ 和 $T_{j-l-1}(1+2r)$,而前者常比后者大.

上面 Kaniel-Paige 的估计，还可以稍许改进，有人考虑如下: 在估计式(27)中 $T_{j-l}(1+2r)$,随 j 增大而增大的速度,依赖于 $r=\dfrac{\lambda_l-\lambda_{l+1}}{\lambda_{l+1}-\lambda_n}$, 如果矩阵 A 的特征值分布如图 4 所示,

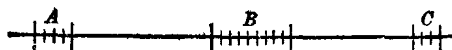

图 4

分成三部分，左端部分 A，中间部分 B，右端部分 C，A 和 C 与中间部分 B 却分隔得较远.

区间 A 有特征值 $\lambda_1, \lambda_2, \cdots, \lambda_{r_1}$,

区间 B 有特征值 $\lambda_{r_1+1}, \lambda_{r_1+2}, \cdots, \lambda_{r_2}$,

区间 C 有特征值 $\lambda_{r_2+1}, \cdots, \lambda_n$,

此时取
$$m(\lambda)=\frac{\prod_{k=1}^{r_1}(\lambda-\lambda_k)}{(\lambda-\lambda_l)}\prod_{k=r_2+1}^{n}(\lambda-\lambda_k)g(\lambda),$$

$$g(\lambda)=T_{j-(n-r_2)-r_1}\left(\frac{2\lambda-\lambda_{r_1+1}-\lambda_{r_2}}{\lambda_{r_2}-\lambda_{r_1+1}}\right),$$

$$g(\lambda_l)^2=T_{j-(n-r_2)-r_1}\left(\frac{2\lambda_l-\lambda_{r_1+1}-\lambda_{r_2}}{\lambda_{r_2}-\lambda_{r_1+1}}\right)^2$$
$$=T_{j-(n-r_2)-r_1}(1+2\gamma_{r_1r_2})^2,$$

$$\gamma_{r_1r_2}=\frac{\lambda_{r_1+1}-\lambda_l}{\lambda_{r_2}-\lambda_{r_1+1}},$$

这样可使 $\gamma_{r_1r_2}$ 比较大，从而会获得更好的估计.

现在开始介绍 Saad 的结果；Paige 为了获得 $\mu_l-\lambda_l$ 的估计考虑的是 $\boldsymbol{s}\in\{\boldsymbol{q}_1, A\boldsymbol{q}_1, \cdots, A^{j-1}\boldsymbol{q}_1\}$，并且 $\boldsymbol{s}$ 正交于特征向量 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_{l-1}$. Saad 考虑的是 $\boldsymbol{s}\in\{\boldsymbol{q}_1, A\boldsymbol{q}_1, \cdots, A^{j-1}\boldsymbol{q}_1\}$，$\boldsymbol{s}$ 正交于 Ritz 向量 $\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_{l-1}$.

定理 4.20 对任意 $s\in\{\boldsymbol{q}_1, A\boldsymbol{q}_1, \cdots, A^{j-1}\boldsymbol{q}_1\}$，如果 s 正交于 Ritz 向量 $\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_{l-1}$，则
$$\mu_l-\lambda_l\leqslant\rho(s, A-\lambda_l I).$$

证明 只要证明
$$\mu_l\leqslant\rho(\boldsymbol{s}, A),$$

因为 $s \in \{q_1, Aq_1, \cdots, A^{j-1}q_1\}$，故 $s \in \{z_1, z_2, \cdots, z_j\}$，因此

$$s = \sum_{k=1}^{j} \beta_k z_k,$$

但 $s^* z_k = 0 (k=1, 2, \cdots, l-1)$，故

$$s = \sum_{k=l}^{j} \beta_k z_k = Q_j \sum_{k=l}^{j} \beta_k s_k,$$

$$\begin{aligned}\rho(s, A) &= (As, s)/(s, s) \\ &= \left(Q_j^* A Q_j \sum_{k=l}^{j} \beta_k s_k, \sum_{k=l}^{j} \beta_k s_k\right) \Big/ \left(\sum_{k=l}^{j} \beta_k s_k, \sum_{k=l}^{j} \beta_k s_k\right) \\ &= (T_j h, h)/(h, h) = \rho(h, T_j),\end{aligned}$$

这里 $h = \sum_{k=l}^{j} \beta_k s_k$，是正交于 T_j 的特征向量 s_1、s_2、…、s_{l-1} 的 j 维向量，因此 $\rho(h, T_j) \geqslant \mu_l$，即

$$\mu_l \leqslant \rho(s, A)\text{、证毕.}$$

从定理 4.20 可以看出 Saad 结果会简单一些。

定理 4.21 对任意自然数 $l \leqslant j \leqslant n$，

$$q_1 = \sum_{k=1}^{n} \alpha_k y_k, \quad \alpha_l \neq 0,$$

如果 $\mu_{l-1} \leqslant \lambda_{l+1}$，则

$$\mu_l - \lambda_l \leqslant (\lambda_n - \lambda_l) \left(\sum_{k=l+1}^{n} \alpha_k^2 \Big/ \alpha_l^2\right) \times \left(\prod_{k=1}^{l-1} \frac{\mu_k - \lambda_n}{\mu_k - \lambda_l} \Big/ T_{j-l}(1+2r)\right)^2, \tag{28}$$

这里

$$r = \frac{\lambda_l - \lambda_{l+1}}{\lambda_{l+1} - \lambda_n}.$$

证明 利用定理 4.18，对任意 $m(\lambda) \in P_{j-1}$，有

$$\rho(m(A)q_1, A - \lambda_l I) \leqslant (\lambda_n - \lambda_l) \left(\sum_{k=l+1}^{n} \alpha_k^2 \Big/ \alpha_l^2\right) \frac{\|m(A)h\|^2}{m(\alpha_l)^2},$$

$s = m(A)q_1 \in \{q_1, Aq_1, \cdots, A^{j-1}q_1\}$，由 Ritz 向量的性质 4 知，要 $s^* z_k = 0 (k=1, 2, \cdots, l-1)$，可取

$$m(\lambda)=(\lambda-\mu_1)(\lambda-\mu_2)\cdots(\lambda-\mu_{l-1})g(\lambda),$$

其中 $g(\lambda)$ 是任意 $j-l$ 次多项式，对于这样的 $m(\lambda)$，利用定理 4.20，有

$$\begin{aligned}\mu_l-\lambda_l&\leqslant\rho(\boldsymbol{s},\ A-\lambda_l I)\\&\leqslant(\lambda_n-\lambda_l)\left(\sum_{k=l+1}^{n}\alpha_k^2\Big/\alpha_l^2\right)\frac{\|m(A)\boldsymbol{h}\|^2}{m(\alpha_l)^2},\end{aligned}$$

但

$$\begin{aligned}\|m(A)\boldsymbol{h}\|&\leqslant\max_{l+1\leqslant i\leqslant n}|(\lambda_i-\mu_1)(\lambda_i-\mu_2)\cdots\\&\quad\cdot(\lambda_i-\mu_{l-1})|\max_{l+1\leqslant i\leqslant n}|g(\lambda_i)|\\&=(\lambda_n-\mu_1)(\lambda_n-\mu_2)\cdots(\lambda_n-\mu_{l-1})\max_{l+1\leqslant i\leqslant n}|g(\lambda_i)|,\end{aligned}$$

特别取
$$g(\lambda)=T_{j-l}\left(\frac{2\lambda-\lambda_{l+1}-\lambda_n}{\lambda_n-\lambda_{l+1}}\right),$$

即得(28). 证毕.

对于 Saad 的结果，也可以取

$$\begin{aligned}m(\lambda)=&\frac{\prod\limits_{k=1}^{r_1}(\lambda-\lambda_k)}{(\lambda-\lambda_l)}\\&\times\prod_{k=r_2+1}^{n}(\lambda-\lambda_k)T_{j-(n-r_2)-r_1}\left(\frac{2\lambda-\lambda_{r_1+1}-\lambda_{r_2}}{\lambda_{r_2}-\lambda_{r_1+1}}\right)\end{aligned}$$

来改善 $\mu_l-\lambda_l$ 的估计.

§4 在有限位精度运算下的 Lanczos 算法

在有限位精度运算下的 Lanczos 算法实施时，会产生舍入误差，它的影响破坏了 $\boldsymbol{q}_l$ 与 $\boldsymbol{q}_k$ 之间的正交性，于是当 j 适当大时，$I-Q_j^*Q_j$ 不再是零矩阵，甚至 $\|I-Q_j^*Q_j\|$ 不是一个小量. 与前述无限位精度运算下的结果相比，带来了复杂性.

例如，对于

$$\beta_{l-1}=\|\boldsymbol{r}_{l-1}\|,$$

$$\boldsymbol{r}_l=A\boldsymbol{q}_l-\alpha_l\boldsymbol{q}_l-\beta_{l-1}\boldsymbol{q}_{l-1},$$

在无限位精度下：

$$\beta_{l-1}=(A\boldsymbol{q}_l,\ \boldsymbol{q}_{l-1}).$$

但在有限位精度下$\|\boldsymbol{r}_{l-1}\|$，与$(A\boldsymbol{q}_l,\ \boldsymbol{q}_{l-1})$就不一定相同.

在无限位精度下

$$\alpha_l=(A\boldsymbol{q}_l,\ \boldsymbol{q}_l)=(A\boldsymbol{q}_l-\beta_{l-1}\boldsymbol{q}_{l-1},\ \boldsymbol{q}_l),$$

但在有限位精度下$(A\boldsymbol{q}_l,\ \boldsymbol{q}_l)$与$(A\boldsymbol{q}_l-\beta_{l-1}\boldsymbol{q}_{l-1},\ \boldsymbol{q}_l)$也不一定相同.

从这两个例子说明，对于每一步计算中，α_l，与β_{l-1}按不同公式计算是有差别的．为了下面讨论方便，我们规定，Lanczos 算法如下：

1．取向量$\boldsymbol{q}_1$，$\|\boldsymbol{q}_1\|=1$，造$\boldsymbol{u}_1=A\boldsymbol{q}_1$，$1\to l$；

2．$\alpha_l=\boldsymbol{q}_l^*\boldsymbol{u}_l$；

3．$\boldsymbol{r}_l=\boldsymbol{u}_l-\alpha_l\boldsymbol{q}_l$；

4．$\beta_l=\|\boldsymbol{r}_l\|$；

5．若β_l充分小，转入计算T_l的特征值，特征向量．否则$\boldsymbol{q}_{l+1}=\boldsymbol{r}_l/\beta_l$；

6．$\boldsymbol{u}_{l+1}=A\boldsymbol{q}_{l+1}-\beta_l\boldsymbol{q}_l$，$l+1\to l$转 2.

这样算法，可使矩阵T_j，保持对称，使得

$$(\boldsymbol{r}_l,\ \boldsymbol{q}_l)=0$$

在工作精度下成立．但是$(\boldsymbol{r}_l,\ \boldsymbol{q}_{l-1})$不一定是零.

在这样的算法下，关系式

$$AQ_j-Q_jT_j=\beta_j\boldsymbol{q}_{j+1}\boldsymbol{e}_j^*$$

也不一定成立了．但可以知道成立

$$AQ_j-Q_jT_j=\beta_j\boldsymbol{q}_{j+1}\boldsymbol{e}_j^*+F_j,\qquad(29)$$

这里 F_j 是由舍入误差引起的 $n\times j$ 矩阵。

从实际计算的经验知

$$\|F_j\|\leqslant\varepsilon\|A\|,$$

这里 ε 是计算机的工作精度. 在 C. Paige 1976 的文章 [27] 中，指出

$$\|F_j\|_F\leqslant(7+\alpha)\sqrt{j}\ \varepsilon\|A\|$$

一定成立. 这里 α 是跟计算 Ax 有关的误差. 如果 A 的每行最多有 m 个非零元，将 A 的元素都取绝对值后的矩阵记为 $|A|$，$\||A|\|/\|A\|=\nu$，则

$$\alpha\leqslant m\nu.$$

关于这一结果，我们不在此介绍了. 我们的兴趣，仍然是估计 T_j 的特征值与 A 的特征值之间的差，也即要回答 T_j 的特征值，是否可以作为 A 的特征值的近似值.

因为现在 Q_j 的列不再正交了，因此我们可以利用 W. Kahan 定理 4.10，得到

$$|\lambda_{i_k}-\mu_k|\leqslant\sqrt{2}\,\|R\|/\sigma_1,$$

这里 $R=\beta_j q_{j+1}e_j^*+F_j$，来估计 μ_k 作为 A 的特征值的近似值. 可是计算经验表明，当 $\boldsymbol{q}_1$、$\boldsymbol{q}_2$、…、$\boldsymbol{q}_j$，j 增大时，不但失去正交性，而且很快就线性相关了. 因此在 j 不太大时，可以用上述估计式. 当 j 较大时，$\sigma_1(Q_j)=0$，就不能使用这个估计式了. 为此我们转而研究(7)型的估计式.

假如 T_j 的特征值 μ_k 所对应的特征向量为 $\boldsymbol{s}_k$，$\|\boldsymbol{s}_k\|=1$，从(29)式有

$$AQ_j\boldsymbol{s}_k-\mu_k Q_j\boldsymbol{s}_k=\beta_j s_{jk}\boldsymbol{q}_{j+1}+F_j\boldsymbol{s}_k,$$

仍记 $\boldsymbol{z}_k=Q_j\boldsymbol{s}_k$ 为 Ritz 向量，于是有

$$A\boldsymbol{z}_k-\mu_k\boldsymbol{z}_k=\beta_j s_{jk}\boldsymbol{q}_{j+1}+F_j\boldsymbol{s}_k,$$

于是利用(7)有 A 的特征值 λ，

$$|\lambda-\mu_k|\leqslant\|\boldsymbol{r}\|/\|\boldsymbol{z}_k\|,$$

这里 $\boldsymbol{r}=\beta_j s_{jk}\boldsymbol{q}_{j+1}+F_j\boldsymbol{s}_k$, 有

$$\|\boldsymbol{r}\|\leqslant\beta_{jk}+\|F_j\|.$$

因为$\|F_j\|$是一个小量，因此只要 β_{jk} 充分小，可使$\|\boldsymbol{r}\|$是一个小量．留下的问题是$\|\boldsymbol{z}_k\|$是否会很小.

尽管$\|\boldsymbol{s}_k\|=1$, 但 Q_j 的列不正交，甚至会线性相关，因此 $Q_j\boldsymbol{s}_k$ 的范数可能会很小.

这样使用 Lanczos 算法，面临二个问题:

1. β_{jk} 是否会很小，什么时候它才很小.

2. $\|z_k\|$是否能保持在1附近.

1971 年 C. Paige 的博士论文 [13] 回答了这二个问题. 他的结论是

1. $\boldsymbol{z}_k$ 与 $\boldsymbol{q}_{j+1}$ 失去正交性 $\rightleftarrows$ β_{jk} 很小;

2. 当 $\min\limits_{i\neq k}|\mu_i-\mu_k|$不很小时，$\|\boldsymbol{z}_k\|$ 不会很小.

下面来介绍这些结果，记

$$I-Q_j^*Q_j=C_j^*+\Delta_j+C_j,$$

这里C_j是严格上三角形，Δ_j是对角阵，在我们规定的Lanczos算法下，$(\tilde{\boldsymbol{q}}_l,\ \boldsymbol{q}_{l+1})=0$, 因此我们有$C_j$的第1条超对角上的元素都为0. Δ_j 的对角元反映 $\|\boldsymbol{q}_l\|=1$ 的误差，因此一般说来也是很小的.

记 $\boldsymbol{s}_1$、$\boldsymbol{s}_2$、…、$\boldsymbol{s}_j$ 是 T_j 正确的单位特征向量，对应的特征值为 μ_1、μ_2、…、μ_j, 因此矩阵

$$S=[\boldsymbol{s}_1,\ \boldsymbol{s}_2,\ \cdots,\ \boldsymbol{s}_j]$$

是一个正交阵. 矩阵 $F_j^*Q_j-Q_j^*F_j$ 是一个斜对称矩阵，因此它的对角元为0, 故可记为

$$F_j^*Q_j-Q_j^*F_j=K-K^*,$$

K 是严格上三角阵，由于 $\|F_j\| \sim \varepsilon\|A\|$，因此 K 一般也不会太大，记 $s_i^*(K-K^*)s_k = l_{ik}^{(j)}$.

$$\Delta_j T_j - T_j \Delta_j$$

也是一个斜对称矩阵，记为 $N-N^*$，N 是严格上三角阵，因为 Δ_j 的元素是很小的量，因此 N 的元素也是小量.

矩阵 $C_j T_j - T_j C_j$ 是一个对角元为 0 的上三角阵. 记矩阵

$$G = S^*(K+N)S = (r_{ik}^{(j)}).$$

定理 4.22 设 $q_l^* q_{l+1} = 0$, $l=1, 2, \cdots, j$, Ritz 向量 $z_i = Q_j s_i$, $i=1, 2, \cdots, j$, 则

1. $z_i^* q_{j+1} = r_{ii}^{(j)} / \beta_j s_{ji}$, $i=1, 2, \cdots, j$; (30)
2. 当 $i \neq k$ 时

$$(\mu_i - \mu_k) z_i^* z_k = r_{ii}^{(j)} \frac{\beta_j s_{jk}}{\beta_j s_{ji}} - r_{kk}^{(j)} \frac{\beta_j s_{ji}}{\beta_j s_{jk}} - l_{ik}^{(j)}. \tag{31}$$

证明 由(29)式

$$AQ_j - Q_j T_j = \beta_j q_{j+1} e_j^* + F_j,$$

得

$$Q_j^* A Q_j - Q_j^* Q_j T_j = \beta_j Q_j^* q_{j+1} e_j^* + Q_j^* F_j,$$

两边取转置

$$Q_j^* A Q_j - T_j Q_j^* Q_j = \beta_j e_j q_{j+1}^* Q_j + F_j^* Q_j,$$

两式相减得

$$T_j Q_j^* Q_j - Q_j^* Q_j T_j = \beta_j Q_j^* q_{j+1} e_j^* - \beta_j e_j q_{j+1}^* Q_j + Q_j^* F_j - F_j^* Q_j, \tag{32}$$

即

$$(I - Q_j^* Q_j) T_j - T_j (I - Q_j^* Q_j)$$
$$= \beta_j Q_j^* q_{j+1} e_j^* - \beta_j e_j q_{j+1}^* Q_j + K^* - K, \tag{33}$$

等式左边为

$$(C_j^*+\Delta_j+C_j)T_j-T_j(C_j^*+\Delta_j+C_j)$$
$$=C_jT_j-T_jC_j+\Delta_jT_j-T_j\Delta_j+C_j^*T_j-T_jC_j^*$$
$$=C_jT_j-T_jC_j+N-N^*+C_j^*T_j-T_jC_j^*,$$

于是比较(33)式等式两边的上三角部分，得到

$$C_jT_j-T_jC_j+N=\beta_jQ_j^*\boldsymbol{q}_{j+1}\boldsymbol{e}_j^*-K,$$
$$\beta_jQ_j^*\boldsymbol{q}_{j+1}\boldsymbol{e}_j^*=C_jT_j-T_jC_j+N+K,$$

两边左乘 $\boldsymbol{s}_i^*$，右乘 $\boldsymbol{s}_i$ 得到

$$\beta_js_{ji}\boldsymbol{z}_i^*\boldsymbol{q}_{j+1}=\mu_i\boldsymbol{s}_i^*C_j\boldsymbol{s}_i-\mu_i\boldsymbol{s}_i^*C_j\boldsymbol{s}_i+r_{ii}^{(j)}=r_{ii}^{(j)},$$

即 $$\boldsymbol{z}_i^*\boldsymbol{q}_{j+1}=r_{ii}^{(j)}/\beta_js_{ji}\quad(j=1,\ 2,\ \cdots,\ j),$$

此即第一个结论.

再在(32)两边，左乘 $\boldsymbol{s}_i^*$，右乘 $\boldsymbol{s}_k$

$$(\mu_i-\mu_k)\boldsymbol{z}_i^*\boldsymbol{z}_k=\beta_js_{jk}\boldsymbol{z}_i^*\boldsymbol{q}_{j+1}-\beta_js_{ji}\boldsymbol{q}_{j+1}^*\boldsymbol{z}_k$$
$$+\boldsymbol{s}_i^*(Q^*F_j-F_j^*Q_j)\boldsymbol{s}_k$$
$$=\beta_js_{jk}r_{ii}^{(j)}/\beta_js_{ji}-\beta_js_{ji}r_{kk}^{(j)}/\beta_js_{jk}-l_{ik}^{(j)},$$

即为第二个结论. 证毕.

从定理 4.22 知道，当 $\boldsymbol{z}_i$ 与 $\boldsymbol{q}_{j+1}$ 正交性好的时候，即 $\boldsymbol{z}_i^*\boldsymbol{q}_{j+1}\sim 0$ 时，β_{ji} 不能很小，反之当 β_{ji} 很小时 $\boldsymbol{z}_i$、$\boldsymbol{q}_{j+1}$ 正交性就差.

从(31)可以知道当 $\mu_i\neq\mu_k$ 时，要 $\boldsymbol{z}_i^*\boldsymbol{z}_k$ 正交性好，β_js_{jk} 与 β_js_{ji} 的数量级应该差不多，如果两者数量级差别很大，就使 $|\boldsymbol{z}_i^*\boldsymbol{z}_k|$ 不是很小.

下面的定理，给出 $\|\boldsymbol{z}_k\|$ 的估计.

定理 4.23 在前述假定下，如记

$$\nu_k=\min_{i\neq k}|\mu_i-\mu_k|,\quad 若\ \nu_k\neq 0,$$
$$\zeta_k=j(j-1)\gamma/\nu_k,$$
$$\gamma=\max_{s,l}|r_{ss}^{(l)}|,$$

则

$$\left|1-\|\boldsymbol{z}_k\|\right| \leqslant \|\Delta_j\|+\zeta_k. \tag{34}$$

证明 $1-\|\boldsymbol{z}_k\|^2=1-\boldsymbol{z}_k^*\boldsymbol{z}_k=1-\boldsymbol{s}_k^*Q_j^*Q_j\boldsymbol{s}_k$

$$=\boldsymbol{s}_k^*(I-Q_j^*Q_j)\boldsymbol{s}_k=\boldsymbol{s}_k^*(C_j^*+\Delta_j+C_j)\boldsymbol{s}_k,$$

因为

$$\|\boldsymbol{s}_k^*C_j\boldsymbol{s}_k\|=|\boldsymbol{s}_k^*C_j^*\boldsymbol{s}_k\|,$$

故

$$\left|1-\|\boldsymbol{z}_k\|^2\right| \leqslant \|\Delta_j\|+2\|\boldsymbol{s}_k^*C_j\boldsymbol{s}_k\|.$$

由此可知：要获得定理 4.23 的结果(34)式，主要的任务是估计$\|\boldsymbol{s}_k^*C_j\boldsymbol{s}_k\|$，为此应用定理 4.22 结果给出以下二个引理.

引理 4.1 记 T_i, $i<j$ 的标准正交特征向量组为 $\boldsymbol{s}_1^{(i)}$, $\boldsymbol{s}_2^{(i)}$, $\cdots$, $\boldsymbol{s}_i^{(i)}$，正交阵 $S_i=[\boldsymbol{s}_1^{(i)}, \boldsymbol{s}_2^{(i)}, \cdots, \boldsymbol{s}_i^{(i)}]$，又记 $j\times i$ 长方阵

$$S_i'=\begin{pmatrix} S_i \\ 0 \end{pmatrix}$$

按照定理 4.22 所用的记号，记向量

$$\boldsymbol{v}_i=\left(\frac{\gamma_{1,1}^{(i)}}{\beta_i s_{i,1}^{(i)}}, \frac{\gamma_{2,2}^{(i)}}{\beta_i s_{i,2}^{(i)}}, \cdots, \frac{\gamma_{i,i}^{(i)}}{\beta_i s_{i,i}^{(i)}}\right)^T, i=1, 2, \cdots, j-1,$$

则

$$\boldsymbol{s}_k^*C_j\boldsymbol{s}_k=-\sum_{l=1}^{j-1}(\boldsymbol{s}_k^*S_l'\boldsymbol{v}_l)\, s_{l+1,k},$$

这里 $s_{l+1,k}$ 是 T_j 的特征向量 $\boldsymbol{s}_k$ 的第 $l+1$ 个分量.

证明

$$-C_j=\begin{pmatrix} 0 & \boldsymbol{q}_1^*\boldsymbol{q}_2 & \boldsymbol{q}_1^*\boldsymbol{q}_3 & \cdots & \boldsymbol{q}_1^*\boldsymbol{q}_j \\ & 0 & \boldsymbol{q}_2^*\boldsymbol{q}_3 & \cdots & \boldsymbol{q}_2^*\boldsymbol{q}_j \\ & & 0 & \ddots & \vdots \\ & & & \ddots & \boldsymbol{q}_{j-1}^*\boldsymbol{q}_j \\ & & & & 0 \end{pmatrix},$$

所以 $-C_j$ 的第 $i+1$ 列为

$$\begin{pmatrix} Q_i^*\boldsymbol{q}_{i+1} \\ 0 \end{pmatrix}=\begin{pmatrix} S_iS_i^*Q_i^*\boldsymbol{q}_{i+1} \\ 0 \end{pmatrix},$$

若记 $Z_i=Q_iS_i$，它是 T_i 的 Ritz 向量组成的矩阵，这样 $-C_j$ 的第 $i+1$ 列，又可写成

$$\begin{pmatrix} S_i Z_i^* \boldsymbol{q}_{i+1} \\ 0 \end{pmatrix},$$

对于 $Z_i^*\boldsymbol{q}_{i+1}$ 可以应用定理 4.22，有

$$Z_i^*\boldsymbol{q}_{i+1}=\left(\frac{r_{1,1}^{(i)}}{\beta_i s_{i,1}^{(i)}},\ \frac{r_{2,2}^{(i)}}{\beta_i s_{i,2}^{(i)}},\ \cdots,\ \frac{r_{i,i}^{(i)}}{\beta_i s_{i,i}^{(i)}}\right)^T=\boldsymbol{v}_i,$$

于是 $-C_j$ 的第 $i+1$ 列为

$$\begin{pmatrix} S_i \boldsymbol{v}_i \\ 0 \end{pmatrix}=S_i'\boldsymbol{v}_i,$$

因此 $$\boldsymbol{s}_k^* C_j \boldsymbol{s}_k=-\sum_{l=1}^{j-1}(\boldsymbol{s}_k^* S_l' \boldsymbol{v}_l)s_{l+1,k}.$$ 证毕.

现在需要估计 $\boldsymbol{s}_k^* S_l'$.

引理 4.2 若记 T_j 的特征值为 $\mu_1,\ \mu_2,\ \cdots,\ \mu_j$，$T_l(l<j)$ 的特征值为 $\mu_1^{(l)}$、$\mu_2^{(l)}$、$\cdots$、$\mu_l^{(l)}$，都是按由小到大次序排列. 记

$$\omega_{l,k}=\begin{cases}\sum\limits_{p=1}^{l} r_{p,p}^{(l)}\Big/(\mu_k-\mu_p^{(l)}), & \text{当 } \mu_k\neq\mu_p^{(l)},\ p=1,\ 2,\ \cdots,\ l, \\ 0, & \text{当 } \mu_k \text{ 与 } \mu_p^{(l)}(p=1,\ 2,\ \cdots,\ l) \text{ 中某一个相同时}.\end{cases}$$

则
$$\boldsymbol{s}_k^* C_j \boldsymbol{s}_k=-\sum_{l=1}^{j-1} s_{l+1,k}^2\omega_{l,k}. \tag{35}$$

证明 S_l' 中的第 p 列为

$$\begin{pmatrix} \boldsymbol{s}_p^{(l)} \\ 0 \end{pmatrix},$$

$\boldsymbol{s}_p^{(l)}$ 是 T_l 的第 p 个特征向量，对应的特征值为 $\mu_p^{(l)}$，

$$T_j\begin{pmatrix} \boldsymbol{s}_p^{(l)} \\ 0 \end{pmatrix}-\mu_p^{(l)}\begin{pmatrix} \boldsymbol{s}_p^{(l)} \\ 0 \end{pmatrix}=\begin{pmatrix} 0 \\ \beta_l s_{l,p}^{(l)} \\ 0 \end{pmatrix}\text{第 } l+1 \text{ 行},$$

两边左乘 s_k^* 得

$$\mu_k \boldsymbol{s}_k^* \begin{pmatrix} \boldsymbol{s}_p^{(l)} \\ 0 \end{pmatrix} - \mu_p^{(l)} \boldsymbol{s}_k^* \begin{pmatrix} \boldsymbol{s}_p^{(l)} \\ 0 \end{pmatrix} = \beta_l s_{l,p}^{(l)} s_{l+1,k},$$

或
$$(\mu_k - \mu_p^{(l)}) \boldsymbol{s}_k^* \begin{pmatrix} \boldsymbol{s}_p^{(l)} \\ 0 \end{pmatrix} = \beta_l s_{l,p}^{(l)} s_{l+1,k},$$

若 $\mu_k - \mu_p^{(l)} \neq 0$, 则

$$\boldsymbol{s}_k^* \begin{pmatrix} \boldsymbol{s}_p^{(l)} \\ 0 \end{pmatrix} = \frac{\beta_l s_{l,p}^{(l)} s_{l+1,k}}{(\mu_k - \mu_p^{(l)})},$$

如果对某个 p, 有 $\mu_k - \mu_p^{(l)} = 0$, 那么必须有

$$\beta_l s_{l,p}^{(l)} s_{l+1,k} = 0,$$

但因为 $l<j$, $\beta_l \neq 0$, $s_{l,p}^{(l)}$ 是 $\boldsymbol{s}_p^{(l)}$ 的最后一个分量，在第 2 章已经指出，它不为 0, 因此必须 $s_{l+1,k}=0$. 这样如果我们记某些自然数的集合

$$m = \{l \mid \{1, 2, \cdots, j-1\} \cap \{\mu_k \neq \mu_r^{(l)}, r=1, 2, \cdots, l\},$$

则有

$$\begin{aligned}
\boldsymbol{s}_k^* C_j \boldsymbol{s}_k &= -\sum_{l \in m} (\boldsymbol{s}_k^* S_l' \boldsymbol{v}_l) s_{l+1,k} \\
&= -\sum_{l \in m} (\boldsymbol{s}_k^* \boldsymbol{s}_1^{(l)}, \boldsymbol{s}_k^* \boldsymbol{s}_2^{(l)}, \cdots, \boldsymbol{s}_k^* \boldsymbol{s}_l^{(l)}) \boldsymbol{v}_l s_{l+1,k} \\
&= -\sum_{l \in m} \beta_l s_{l+1,k}^2 \left(\frac{s_{l,1}^{(l)}}{\mu_k - \mu_1^{(l)}}, \frac{s_{l,2}^{(l)}}{\mu_k - \mu_2^{(l)}}, \cdots, \frac{s_{l,l}^{(l)}}{\mu_k - \mu_l^{(l)}} \right) \boldsymbol{v}_l \\
&= -\sum_{l \in m} \beta_l s_{l+1,k}^2 \sum_{p=1}^{l} \frac{s_{l,p}^{(l)} r_{p,p}^{(l)}}{(\mu_k - \mu_p^{(l)}) \beta_l s_{l,p}^{(l)}} \\
&= -\sum_{l \in m} s_{l+1,k}^2 \sum_{p=1}^{l} \frac{r_{p,p}^{(l)}}{\mu_k - \mu_p^{(l)}},
\end{aligned}$$

因此若记
$$\omega_{l,k} = \begin{cases} \displaystyle\sum_{p=1}^{l} \frac{r_{p,p}^{(l)}}{\mu_k - \mu_p^{(l)}}, & l \in m, \\ 0, & l \overline{\in} m, \end{cases}$$

(这就是本引理结论中的量 $\omega_{l,k}$)就有

$$s_k^* C_j s_k = -\sum_{l=1}^{j-1} s_{l+1,k}^2 \omega_{l,k}. \quad \text{证毕.}$$

用引理 4.2，如果我们假定

$$\tilde{\nu}_k = \min_{\substack{\mu_r^{(l)} \neq \mu_k \\ l=1,2,\cdots,j-1 \\ r=1,2,\cdots,l}} |\mu_k - \mu_r^{(l)}|,$$

那么我们可以很容易获得类似于定理的结果. 但是 $\tilde{\nu}_k$ 这个量可能很小很小. 因此这样获得的结果是粗糙的，为了获得更好的结果，将 $\tilde{\nu}_k$ 换成 ν_k. 下面要想法把 $\mu_k-\mu_r^{(l)}$ 换成 $\mu_k-\mu_i$.

现在再来继续证明定理 4.23，考虑量 $s_{l+1,k}^2$，由第 2 章定理 2.9 公式(15)，

$$s_{l+1,k}^2 = \chi_{1,l}(\mu_k)\chi_{l+2,j}(\mu_k)/\chi'_{1,j}(\mu_k),$$

$\chi_{1,l}(\lambda)$ 是 T_l 的特征多项式，它的根为 $\mu_1^{(l)}$、$\mu_2^{(l)}$、$\cdots$、$\mu_l^{(l)}$，$\chi_{l+2,j}(\lambda)$ 有 $j-l-1$ 个根. 将 $\chi_{1,l}(\lambda)$ 和 $\chi_{l+2,j}(\lambda)$ 的全部 $j-1$ 个根，按从小到大次序排列成 ω_1、ω_2、$\cdots$、ω_{k-1}、ω_{k+1}、$\cdots$、ω_j，因为 $\chi'_{1,j}(\mu_k) = \prod\limits_{\substack{i=1 \\ i\neq k}}^{j}(\mu_k-\mu_i)$，因此

$$s_{l+1,k}^2 = \prod_{i=1}^{k-1} \frac{\mu_k-\omega_i}{\mu_k-\mu_i} \prod_{i=k+1}^{j} \frac{\mu_k-\omega_i}{\mu_k-\mu_i}.$$

因为 $\omega_1, \omega_2, \cdots, \omega_{k-1}, \omega_{k+1}, \cdots, \omega_j$ 是矩阵 T_j 划去第 $l+1$ 行、第 $l+1$ 列后的矩阵 B_{l+1} 的特征值，利用极大极小原理可知

$$\left.\begin{aligned}
&\mu_s \leqslant \omega_s \quad (s=1, 2, \cdots, k-1), \\
&\mu_k \leqslant \omega_{k+1}, \\
&\mu_{j-s} \geqslant \omega_{j-s} \quad (s=0, 1, \cdots, j-k-1), \\
&\mu_k \geqslant \omega_{k-1}.
\end{aligned}\right\} \tag{36}$$

实际上，如果对于第 $l+1$ 个分量为 0 的向量 $\boldsymbol{x}$，有

$$\rho(\boldsymbol{x}, T_j) = \rho(\boldsymbol{x}, B_{l+1}),$$

若记 $\widetilde{V}_s$ 是 j 维向量的一个 s 维子空间，这些向量的第 $l+1$ 个分量为 0. 于是

$$\begin{aligned}\mu_s&=\min_{V_s}\max_{x\in V_s}\rho(\boldsymbol{x},\ T_j)\leqslant\min_{\widetilde{V}_s}\max_{x\in\widetilde{V}_s}\rho(\boldsymbol{x},\ T_j)\\&=\min_{\widetilde{V}_s}\max_{x\in\widetilde{V}_s}\rho(\boldsymbol{x},\ B_{l+1})\\&=B_{l+1}\text{ 的第 } s \text{ 个特征值}\quad(s\leqslant j-1).\end{aligned}$$

同样

$$\begin{aligned}\mu_{j-s+1}&=\max_{V_s}\min_{x\in V_s}\rho(\boldsymbol{x},\ T_j)\geqslant\max_{\widetilde{V}_s}\min_{x\in\widetilde{V}_s}\rho(\boldsymbol{x},\ T_j)\\&=\max_{\widetilde{V}_s}\min_{x\in\widetilde{V}_s}\rho(\boldsymbol{x},\ B_{l+1})\\&=B_{l+1}\text{ 的第 } j-s+1 \text{ 个特征值},\quad s\leqslant j-1.\end{aligned}$$

由(36)式可知

$$\left|\frac{\mu_k-\omega_i}{\mu_k-\mu_i}\right|\leqslant 1\quad(i=1,2,\cdots,k-1,k+1,k+2,\cdots,j).\tag{37}$$

另一方面对每个 $r\leqslant l$, 在 $\omega_1,\ \omega_2,\ \cdots,\ \omega_{k-1},\ \omega_{k+1},\ \cdots,\ \omega_j$ 中必有一个是 $\mu_r^{(l)}$, 记为第 $i(r)$ 个，即 $\mu_r^{(l)}=\omega_{i(r)}$, 并且 $i(r)\neq k$, 于是

$$\begin{aligned}s_{l+1,k}^2&=\prod_{i=1}^{k-1}\frac{\mu_k-\omega_i}{\mu_k-\mu_i}\prod_{i=k+1}^{j}\frac{\mu_k-\omega_i}{\mu_k-\mu_i}\\&=g_{i(r)}\frac{\mu_k-\mu_r^{(l)}}{\mu_k-\mu_{i(r)}}\quad(r=1,\ 2,\ \cdots,\ l),\end{aligned}$$

其中

$$g_{i(r)}=s_{l+1,k}^2\Big/\frac{\mu_k-\omega_{i(r)}}{\mu_k-\mu_{i(r)}},$$

由(37)知 $|g_{i(r)}|\leqslant 1$.

由此

$$\begin{aligned}\boldsymbol{s}_k^*C_j\boldsymbol{s}_k&=-\sum_{l=1}^{j-1}s_{l+1,k}^2\omega_{l,k}\\&=-\sum_{l=1}^{j-1}\sum_{p=1}^{l}g_{i(p)}\frac{\mu_k-\mu_p^{(l)}}{\mu_k-\mu_{i(p)}}\frac{\gamma_{p,p}^{(l)}}{\mu_k-\mu_p^{(l)}}\\&=-\sum_{l=1}^{j-1}\sum_{p=1}^{l}g_{i(p)}\frac{\gamma_{p,p}^{(l)}}{\mu_k-\mu_{i(p)}},\end{aligned}$$

从而

$$|s^*C_js_k| \leqslant \sum_{l=1}^{j-1}\sum_{p=1}^{l}\frac{|r_{p,p}^{(l)}|}{|\mu_k-\mu_{i(p)}|} \leqslant \frac{r}{\nu_k}\sum_{l=1}^{j-1}\sum_{p=1}^{l}1$$

$$\leqslant \frac{j(j-1)}{2}r/\nu_k. \quad 证毕.$$

$r_{s,s}^{(l)}$ 是跟 $\varepsilon\|A\|$ 相当的小量，因此 r 也是跟 $\varepsilon\|A\|$ 相当的小量，故若 ν_k 不是太小，ζ_k 一般也是一个小量．于是从定理 4.22 知 $1-\|z_k\|$，在 ν_k 不是太小的条件下，也是一个小量．

至此我们介绍了 Lanczos 的理论基础．在实际计算时 Lanczos 算法过程中，常常会出现这样的问题：某个 T_j 的二个特征值或更多个特征值，对应逼近的是 A 的同一个特征值．这一现象的出现，是由于 Q_j 的正交性失掉而产生的．在计算中如果碰到这种情况，就要把这些特征值的 Ritz 向量算出来，看看是否平行或几乎平行．如果是平行或几乎平行，那么这些 Ritz 值就对应 A 的同一个特征值．

从上面的分析可知，μ_k 是否接近于 A 的特征值，关键是看量 β_{jk} 是否足够小．但 $\beta_{jk}=|\beta_j s_{jk}|$，$\beta_j$ 是已知的量，s_{jk} 是 T_j 的第 k 个特征向量的最后一个分量．如果能提供一种方法，直接从 T_j 计算得 s_{jk}，对 Lanczos 算法是很有价值的．

下面我们来介绍一种求 T_j 的特征向量的办法，假定特征值 μ_k 已经知道了．

$$(T_j-\mu_k I)z=0, \tag{38}$$

$z=(\eta_1, \eta_2, \cdots, \eta_j)^T$，令 $\eta_1=1$，理论上可以从(38)的第 1 个方程求得 η_2，再从第二个方程求得 η_3，依此类推可以从(38)的第 $j-1$ 个方程算得第 j 个分量 η_j．但是这样的方法在 Wilkinson 的书[20]中，已经分析过，是不稳定的．原因是我们使用的是近似特征值 $\tilde{\mu}_k$，计算又有舍入误差，因此 η_{j-1} 和 η_j 两个量代入到(38)的第 j 个方程左边，

$$(\alpha_j-\mu_k)\eta_j+\beta_{j-1}\eta_{j-1}$$

不会是 0，如果它是 δ，于是得到方程

$$(T_j-\tilde{\mu}_k I)z=\delta e_j,$$

这样 $$z=\delta\sum_{l=1}^{j}\frac{(e_j,\ s_l)}{\mu_l-\tilde{\mu}_k}s_l=\delta\sum_{l=1}^{j}\frac{s_{jl}}{\mu_l-\tilde{\mu}_k}s_k,$$

如果 $\dfrac{s_{jk}}{\mu_k-\tilde{\mu}_k}$ 的数量级比其余 $\dfrac{s_{jl}}{\mu_l-\tilde{\mu}_k}$ 的数量级要大，那么可以使 z 是 s_k 方向好的近似向量. 否则 z 就不是好的近似特征向量.

同样从 $\eta_j=1$，利用方程(38)朝前推，求得 η_{j-1}、η_{j-2}、…、η_1 的方法，也有同样的问题. 此时得到的方程是

$$(T_j-\tilde{\mu}_k I)z=\delta e_1,$$

B. N. Parlett 和 J. K. Reid 在文[28]中，提出如下的方法：对 $(T_j-\tilde{\mu}_k I)z=0$，令 $\eta_1=1$，从第 1 个方程求出 η_2，依次到从第 $l-1$ 个方程求出 η_l.

再令 $\xi_j=1$，从第 j 个方程

$$(\alpha_j-\tilde{\mu}_k)\xi_j+\beta_{j-1}\xi_{j-1}=0$$

求出 ξ_{j-1}，依次到从第$(l+1)$个方程求得 ξ_l.

再将分别从两端出发求得的两组数啮合起来，即令

$$\xi_i=\eta_i\frac{\xi_l}{\eta_l}\quad(i=1,\ 2,\ \cdots,\ l-1),\tag{39}$$

这样得到的向量 $z=(\xi_1,\ \xi_2,\ \cdots,\ \xi_j)^T$，代入第 l 个方程左边有

$$\beta_{l-1}\xi_{l-1}+(\alpha_l-\tilde{\mu}_k)\xi_l+\beta_l\xi_{l+1}=\delta,\tag{40}$$

于是 z 满足 $$(T_j-\tilde{\mu}_k)z=\delta e_l.$$

现在来考虑如何选择指标 l，使得 $(z,\ e_l)$ 尽可能大，也即 z 的分量中最大的一个对应的次序作为 l. 为此从 $\eta_1=1$、η_2、… 到某 η_m，绝对值都逐个不减少，但 $|\eta_{m+1}|<|\eta_m|$，那么

这个m就是 l 的一个候选者. 再从 $\xi_j=1$、ξ_{j-1}、$\cdots$ 逐个计算到 ξ_m, 通过啮合公式(40)求出 $\boldsymbol{z}=(\xi_1, \xi_2, \cdots, \xi_j)^T$ 后, 将它的分量代入(40), 得到 δ, 考察 $\delta/\|\boldsymbol{z}\|$, 如果这是一个小量, 那么 $s=\boldsymbol{z}/\|\boldsymbol{z}\|$ 即为所求的单位特征向量的近似. 否则, 再从 η_{m+1} 开始求第二个极大值 $\eta_{m'}$, 当然 $|\eta_{m'}|$ 要求比 $|\eta_m|$ 大, 以 m' 作为 l 的候选者, 以此类推.

为什么要以判定 $\delta/\|\boldsymbol{z}\|$的大小, 作为一个判据? 利用关系式(29)

$$AQ_j-Q_jT_j=\beta_j\boldsymbol{q}_{j+1}\boldsymbol{e}_j^*+F_j,$$

和
$$(T_j-\tilde{\mu}_kI)\boldsymbol{s}=\delta e_l/\|\boldsymbol{z}\|,$$

因此
$$AQ_j\boldsymbol{s}-Q_jT_j\boldsymbol{s}=\beta_j\boldsymbol{q}_{j+1}\boldsymbol{e}_j^*\boldsymbol{s}+F_j\boldsymbol{s},$$

有
$$AQ_j\boldsymbol{s}-\tilde{\mu}_kQ_j\boldsymbol{s}=\beta_j\boldsymbol{q}_{j+1}\boldsymbol{e}_j^*\boldsymbol{s}+F_j\boldsymbol{s}+\frac{\delta}{\|\boldsymbol{z}\|}Q_j\boldsymbol{e}_l,$$

记 $\boldsymbol{y}=Q_j\boldsymbol{s}$, 又记 s 最后一个分量为 $\tilde{s}_{jk}$, 即有

$$A\boldsymbol{y}-\tilde{\mu}_k\boldsymbol{y}=\beta_j\tilde{s}_{jk}\boldsymbol{q}_{j+1}+F_j\boldsymbol{s}+\frac{\delta}{\|\boldsymbol{z}\|}\boldsymbol{q}_l,$$

$\boldsymbol{y}$ 逼近特征向量的程度、$\tilde{\mu}_k$ 逼近特征值的程度, 都依赖于

$$r=\beta_j\tilde{s}_{jk}\boldsymbol{q}_{j+1}+F_j\boldsymbol{s}+\frac{\delta}{\|\boldsymbol{z}\|}\boldsymbol{q}_l$$

的范数, 而 r 的第三项的范数为 $\delta/\|\boldsymbol{z}\|$, 因此使用 $\delta/\|\boldsymbol{z}\|$作为判据.

利用反迭代法求特征向量, 也是一个稳定的方法, 只是存贮量和计算量比上述方法稍许大一点, 参见[28].

最后再说一下, 使用 Lanczos 算法一般地 j 应该取多大, 根据 Parlett 的经验[10, p.259], 当 $n=10^4$, 取 $j=300$, 可以求出 10 个两端部的特征值和对应的特征向量. 要进一步求更多的特征值和特征向量, 可以继续增大 j, 但这样要遇到已求得

的特征值的重复出现的现象．也可以重新取一个初始单位向量 $\tilde{q}_1$，令它与已经求得的特征向量正交，从 $\tilde{q}_1$ 出发再施行 Lanczos 过程．

适当地修改 Lanczos 过程，防止特征值重复出现的现象，仍然是一个研究课题，参见[29]、[30]。

参 考 文 献

[1] M. R. Hestenes and E. Stiefel, Mathods of conjugate gradients for solving linear systems, NBS J. Res. 49 pp. 409-436, 1952.

[2] J. K. Reid, On the method of conjugate gradients for the solution of large sparse system of linear equations, Proceedings of the Conference on Large Sparse Sets of Linear Equations(Ed. J. K. Reid), Academic Press, pp. 231-254. 1971.

[3] G. W. Stewart, The conver gence of the method of conjugate gradients at isolated extreme Points of the spectrum, Numer. Math. 24: 85-93 (1975).

[4] Dianne P. O'Leary, The block conjugate gradient algorithm and related methods, Linear Algebra and Its Applications 29:293-322, 1980.

[5] 蒋尔雄 轴对称热应力问题的算法和程序，748 会议论文集，1974.

[6] 蒋尔雄 对称正定矩阵 P-条件数的改善，复旦大学学报，自然科学版，Vol. 9, No. 1:1-7, 1964.

[7] J. A. Meigerink and H. A. Van der Vorst, An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix, Mathematics of Computation 31:148-162, 1977.

[8] David S. Kershaw, The incomplete Cholesky-conjugate gradient method for the iterative solution of systems of linear equations, Journal of Computational Physics 26, 43-65, 1978.

[9] 蒋尔雄、高坤敏、吴景琨 线性代数，人民教育出版社 1978.

[10] B. N. Parlett, The Symmetric Eigenvalue Problem, Prentice-Hall, Inc. Englewood Cliffs, N. J. 1980.

[11] C. H. Reinsch, A stable, rational QR algorithm for the computation of the eigenvalues of an Hermitian tridiagonal matrix, Math. of Comp., Vol. 25, No. 115, 1971.

[12] R. C. Thompson and P. McEnteggert, Principal submatrices II, the upper and lower Quatratic inequalities, Linear Algebra and its

Appl., 1:211-243, 1968.

[13] C. C. Paige, The computation of eigenvalues and eigenvectors of very large sparse matrices, Ph. D. thesis, Univ. of London, 1971.

[14] Harry Hochstadt, On some inverse problems in matrix theory, Arch. Math. 18:201-207, 1967.

[15] Ole H. Hald, Inverse eigenvalne problems for Jacobi matrices, Linear Algebra Appl. 14:63-85, 1976.

[16] C. de Boor and G. H. Golub, The numerically stable reconstruction of a Jacobi matrix from spectral data, Linear Algebra and its Appl. 21:245-260, 1978.

[17] 曹维潞　关于一种特殊形状线性方程组的求解问题，计算数学，1978 年第 1 期。

[18] C. C. Paige and Saunders, Solution of Sparse indefinite systems of linear equations, SIAM J. Numer. Anal. 12:617-629, 1975.

[19] B. N. Parlett, A new look at the Lanczos algorithm for solving symmetric systems of linear equations, Linear Algebra and its Applications. 29:323-346, 1980.

[20] J. H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press. Oxford. 1965.

[21] J. H. Wilkinson, Golbal convergence of tridiagonal QR algorithm with origin shifts, J. Linear Algebra and Its Applications. Vol. 1:409-420, 1968.

[22] C. L. Lawson and R. J. Hanson, Solving Least Squares Problems, Prentice-Hall, INC., 1974.

[23] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Res. Nat. Bur. Standards, Sect. B 45:225-280, 1950.

[24] H. R. Schwarz, Tridiagonalization of a symmetric band matrix, Numer. Math. 12:231-241, 1968.

[25] S. Kaniel, Estimates for some computational techniques in linear algebra, Math. Comp. 20:369-378, 1966.

[26] Y. Saad, Error bounds on the interior Rayleigh-Ritz approximations from Krylov subspaces, Soc. Ind. Appl. Math. J. Num. Anal. 17, 1980.

[27] C. Paige, Error analysis of Lanczos algorithm for tridiagonalizing a

symmetric matrix, J. Inst. Math. Appl. 18:341-349, 1976.

[28] B. N. Parlett and J. K. Reid, Tracking the progress of the Lanczos algorithm for large symmetric eigenproblems, IMA Jour Num. Analysis, Vol. 1:135-155, 1981.

[29] Parlett, B. N. and Scott, D. S. The Lanczos algorithm with selective orthogonalization, Math. Comp. 33:217-238, 1979.

[30] Horst D. Simon, The Lanczos algorithm for solving symmetric linear systems, Ph. D. Thesis, University of California, Berkeley, 1982.

[31] 蒋尔雄 对称三对角矩阵带位移的 QL 方法的收敛率，高等学校计算数学学报, Vol. 6, No. 3, 1984.

[32] Jiang Erxiong and Zhang Zhenyue, A new shift of the QL algorithm for irreducible symmetric tridiagonal matrices, Linear Algebra and Its Applications, 将发表。

[General Information]
书名=对称矩阵计算
作者=蒋尔雄
页数=160
SS号=10184133
DX号=
出版日期=1984年11月第1版
出版社=上海科学技术出版社